Background documentation for Day 2

This file contains the slides that were shown by the presenters during Day 2 of the meeting as well the background documentation shared with MPAG members ahead of the meeting.

Tuesday, 31 October 2023			
	Session 5	**Open**	
12:00 – 13:00	Comparative effectiveness in the context of the arrival of new vector control products • Report resolution • New SOPs Background |Background 2 | Presentation	Dr Jan Kolaczinski Unit Head, Vector Control & Insecticide Resistance Mr Dominic Schuler Acting Team Lead Vector Control Product Assessment Dr Geraldine Foster, Technical Officer Vector Control Product Assessment	**For information**
	Session 6	**Open**	
13:00 – 14:00	Update on the "High burden to high impact" (HBHI) approach • Overview and updates on ongoing activities • Evaluation report of four HBHI countries Background | Presentation 1 | Presentation 2	Dr Maru A. Weldedawit Unit Head, High Burden to High Impact Prof. Evelyn Ansah Director, Centre for Malaria Research, University of Health & Allied Sciences, Ghana	**For information**
	Session 7	**Open**	
14:00 – 14:15	Update on the "WHO Guidelines for malaria" • Tafenoquine recommendation • G6PD diagnostics Presentation	Dr Andrea Bosman Unit Head, Diagnostics, Medicines & Resistance Dr Peter Olumese Medical Officer, Diagnostics, Medicines & Resistance	**For information**

WHO/UCN/GMP/MPAG/2023.05

	Session 8	**Open**	
14:45 – 15:45	Update on antimalarial drug resistance in Africa Presentation The Mekong malaria elimination programme Background | Presentation	Dr Charlotte Rasmussen Technical Officer, Diagnostics, Medicines & Resistance Dr Pascal Ringwald Coordinator, Malaria and Neglected Tropical Diseases WPRO	**For information**
15:45 – 16:15	Closing session Presentation	Dr Daniel Ngamije M. Director, Global Malaria Programme	**For information**

Malaria Policy Advisory Group Meeting

30 October–1 November 2023, Geneva, Switzerland
Background document for Session 4

World Health Organization

Data requirements and protocol for determining comparative efficacy of vector control products, with a focus on insecticide-treated nets and indoor residual spraying products. Second edition

Final draft for MPAG

Contents

Abbreviations

AI	active ingredient
CI	confidence interval
FIC	first-in-class
GLP	Good Laboratory Practice
IRS	indoor residual spraying
ITN	insecticide-treated net
LS	Latin square
OR	odds ratio
PBO	piperonyl butoxide
RCT	randomized controlled trial
RR	rate ratio
SIC	second-in-class
WHO	World Health Organization

Glossary

Active comparator
In comparative efficacy trials, the active comparator is the reference product to which the candidate product is compared. Ideally, this should be a first-in-class product, but there may be situations in which a second-in-class product provides a suitable alternative. In comparative efficacy trials, a non-inferiority analysis seeks to determine whether a new candidate product is not worse than an active comparator by more than a predefined acceptable amount (called the non-inferiority margin). See also control.

Blood-feeding inhibition
The proportional reduction of blood feeding in huts with insecticide-treated nets relative to controls with untreated nets. Note: Blood-feeding inhibition may be reported but is not used for decision-making on insecticide-treated net or indoor residual spraying products.

Blood-feeding rate
The proportion of blood-fed female mosquitoes relative to the total number of female mosquitoes found in an experimental hut.

Candidate product
A new product that is thought to have the same entomological effect as a product in an established intervention class, but that has yet to provide the entomological evidence for it to be considered covered by a WHO recommendation or to inform extension of an existing recommendation or the formulation of a new one.

Comparative efficacy assessment
Comparative efficacy assessments in the context of vector control evaluations are entomological trials used to determine the entomological performance of products against an active comparator. This evaluation of products enables direct comparison of the relative performance between products and avoids the issues of data comparability (differences in point estimates) introduced by testing different products at different sites and at different times.

Control
A negative control is a product that resembles the candidate product but does not have the active ingredient. A negative control is used to monitor the quality of the evaluation by ensuring that the observed mortality (or other effect) is due to the active ingredient(s) and not to poor execution of the study (e.g. induced mortality from poor handling of mosquitoes).

Current standard of care
The type of vector control product predominantly used by the national malaria control programme in the country where the study will be implemented or where the intervention is expected to be deployed.

Entomological effect
Entomological effect refers to a product's effect on a disease vector in terms of decreasing vector survival, reducing biting rates, reducing fertility, reducing human–vector contact or reducing the vector's susceptibility to infection or transmission. Products with different biochemical modes of action may have similar entomological effects on target insects; for example, indoor residual spraying formulations with pyrethroids and carbamates differ in their chemical modes of action, yet are considered to have a similar impact on the target insect in areas of insecticide susceptibility.

Equivalence trial
An equivalence trial seeks to demonstrate that the test product is not better or worse than the comparator by more or less than an amount known as the equivalence margin. The null hypothesis is that the treatments differ, and the trial is powered to accept or reject this hypothesis.

First-in-class
First-in-class refers to the first product with a novel entomological effect (e.g. decreasing vector survival, reducing biting rates, reducing fertility, reducing human–vector contact or reducing the vector's susceptibility to infection or transmission), the public health value of which is assessed by the Vector Control Advisory Group based on the demonstration of the product's protective efficacy to reduce or prevent infection and/or disease in humans. Once the public health value of a first-in-class product has been determined, a World Health Organization recommendation will be issued, establishing a new intervention class.

Intervention class
An intervention class is defined as a group of interventions with a similar entomological effect and mechanism by which the effect is derived. For a new intervention class to be established, two independently powered trials with epidemiological end-points must demonstrate a significant reduction in the primary epidemiological end-point. Based on this evidence, a guideline development process is initiated to formulate a World Health Organization recommendation.

Note: An intervention class is disease-specific, meaning that for interventions with different target diseases, epidemiological impact needs to be demonstrated against each vector-borne disease to establish the class and inform development of a disease-specific World Health Organization recommendation via the guideline development process.

Mortality
Any mosquito that cannot stand or fly in a coordinated manner or that shows no movement, usually measured at 24 hours after exposure to an intervention. This post-exposure period may be extended depending on the mode of action of the active ingredient in the intervention under evaluation.

Non-inferiority margin
The non-inferiority margin is the maximum worse difference in vector control product efficacy that is predefined to be tolerable from a public health standpoint. Values for the non-inferiority margin (called "delta") for insecticide-treated nets and indoor residual spraying have been derived from mathematical modelling of possible impact on disease should a substantially worse product become available for use in public health. Non-inferiority margins for other vector control products currently under epidemiological evaluation will need to be defined in due course, once these products have been recommended for use.

Non-inferiority trial
A non-inferiority trial is one type of design for a comparative efficacy study. It seeks to demonstrate that the candidate vector control intervention is not worse than the active comparator by more than a pre-specified amount. This amount is known as the non-inferiority margin. The null hypothesis in non-inferiority trials is that the new intervention is inferior to the standard intervention by more than the non-inferiority margin. The alternative hypothesis is that the new intervention is non-inferior to the standard intervention, i.e. that the observed range of efficacy (95% confidence interval) is within the non-inferiority margin.

Odds ratio
The odds ratio is a measure of how the odds of an outcome differ for two different groups. In the context of the primary end-points of the analyses outlined in this document, the odds ratio is calculated from the odds of a mosquito dying or being sterilized for the candidate product versus the odds for the same outcome for the active comparator product. For the analyses outlined in this document, odds ratio = 1 indicates no difference between the active comparator and candidate product, odds ratio > 1 indicates that the event is more likely in the candidate product, and odds ratio < 1 indicates that the event is more likely in the active comparator product.

Primary end-point
The main outcome to be evaluated upon which the comparative efficacy trial is powered. In the context of comparative efficacy trials for vector control products, the primary end-point is used to make the ultimate decision regarding the non-inferiority of a product and its inclusion under a World Health Organization recommendation for an intervention. The choice of primary end-point will depend on the mode of action of the product.

Public health value
A product has public health value if it has proven protective efficacy to reduce or prevent infection and/or disease in humans.

Rate ratio
Rate ratio is an expression of the frequency with which an event occurs in a defined population in a specified time period relative to the frequency in the comparator population. In the context of the primary end-points of the analyses outlined in this document, when assessing a sterility-inducing product, the number of eggs laid may be assessed. The number of eggs is a count outcome, and it is appropriate to use the rate ratio, which calculates the rate of eggs laid for the candidate product versus the rate for the same outcome for the active comparator product. For the analyses outlined in this document, rate ratio = 1 indicates no difference between the active comparator and candidate product, rate ratio > 1 indicates that the event is more likely in the candidate product, and rate ratio < 1 indicates that the event is more likely in the active comparator.

Residual efficacy
The residual efficacy is the duration for which the entomological effect of a vector control product remains above a defined level.

Second-in-class
Second-in-class refers to products that have demonstrated a non-inferior entomological effect relative to the first-in-class product but have not undergone epidemiological evaluation at the time of evaluation. A second-in-class product may be covered by the same World Health Organization recommendation as the first-in-class product or, in the case of considerable divergence, may require formulation of a new recommendation.

Secondary end-point
Secondary end-points are outcomes measured in a trial that are not considered to be the primary outcomes on which decisions are made. In the context of the analyses outlined in this document, secondary end-points remain important in terms of understanding entomological modes of action and how a product functions, but they are not used to determine the necessary size of the study. Blood-feeding inhibition, for example, is a secondary end-point for indoor residual spraying and insecticide-treated net products.

Standard comparator
A product that is the current standard of care in the country where the study is being conducted or where the intervention is expected to be deployed, and that belongs to another/older class, i.e. currently pyrethroid-only insecticide-treated nets for trials of new insecticide-treated nets and a World Health Organization-prequalified product in common use for studies of new indoor residual spraying products.

Study
In the context of the vector control product evaluations outlined in this document, a study is an individually powered evaluation of a product to demonstrate non-inferiority to the first-in-class product. The World Health Organization currently requires two successful independently powered studies for a product to join an intervention class. If feasible, studies should provide geographical diversity and be conducted in at least two different regions (including at least two of East, Central and West Africa). The terms “study” and “trial” are often used interchangeably.

Superiority trial
A superiority trial seeks to demonstrate that the candidate product is better than the comparator. The null hypothesis in superiority trials is that the new treatment is not better than the standard treatment/placebo, and the trial is powered to reject this hypothesis if the test product is superior by a specified amount.

Executive summary

The World Health Organization (WHO) evaluates vector control products with the aim of providing assurance to its Member States that interventions have proven protective efficacy to reduce or prevent infection and/or disease in humans and that specific products meet quality, safety and efficacy standards. New products for which evidence of disease impact is lacking follow the new intervention pathway to assess their public health value, a process supported by the WHO Vector Control Advisory Group. Concomitantly, the safety, efficacy and quality of the product are assessed by the WHO Prequalification Team for Vector Control Products. Subsequent products in the same intervention class must still go through the prequalification process, but are not required to demonstrate public health value, provided that they are found to be non-inferior to the first-in-class product with respect to the applicable entomological end-point. To demonstrate this, manufacturers must generate comparative efficacy data on the entomological impact of the new product relative to the first-in-class product. These data can be generated through the same experimental hut trials used to generate data for the WHO prequalification assessment dossier.

WHO requires these data as indirect evidence for assessing whether a product provides similar impact to the first-in-class product that generated epidemiological data to inform the development of a WHO recommendation for an existing intervention class. In a comparative entomological efficacy assessment, a product needs to demonstrate:

- non-inferiority to the first-in-class product (active comparator) on the primary end-point(s); and
- superiority over the control or current standard of care (standard comparator) on the primary end-point(s), if applicable.

Evaluation of the comparative efficacy of malaria vector control products is part of the classification of vector control products by the Global Malaria Programme and is directly linked to its guideline development process. For all established intervention classes for which epidemiological data were used to inform development of associated WHO recommendations, comparative entomological data provide the necessary assurance of similar product performance and a ready means to address the increasing diversity of products within broadening classes. In turn, these data inform the WHO guideline development process by validating whether existing WHO recommendations are directly applicable to new products, or by informing the extension of a recommendation or development of a new one. The aim of the comparative analysis of products is to provide a relatively easy and cost-effective means of determining the entomological performance of products against a comparator, using data generated through studies required as part of product evaluation for prequalification. Comparative analysis avoids the difficulties of data comparability introduced by testing different products separately at different sites and at different times. In addition to validating whether an existing WHO recommendation applies to a new product, the comparative efficacy analysis aims at providing additional information to inform procurement decisions/product selection by WHO Member States and their implementing partners under increasingly resource-constrained conditions.

1. Background information and rationale

Since 1 January 2017, the World Health Organization (WHO) has implemented a new process for evaluating vector control products *(1)*. The process seeks to provide enhanced assurance regarding product safety, quality and efficacy (both entomological and epidemiological) to better meet the needs of WHO Member States. The assessment of individual products for their quality, safety and entomological efficacy is overseen by the WHO Prequalification Team for Vector Control Products.

The WHO technical departments, namely the Global Malaria Programme and the Department of Control of Neglected Tropical Diseases, review epidemiological data to assess the public health value of new vector control interventions, which in turn informs the development of WHO recommendations through the guideline development process *(2)*. Assessment of public health value is the mandate of the WHO Vector Control Advisory Group, while the WHO guidelines process is supported by specific guideline development groups and is overseen by the Guidelines Review Committee *(3)*.

The evaluation process for vector control products has, over the last six years, continued to evolve. Where needed, implementation experience has been incorporated into the process. As a part of these efforts, the WHO Global Malaria Programme and Department of Control of Neglected Tropical Diseases, with the support of the Vector Control Advisory Group, reviewed and reduced the overall number of intervention classes. With fewer intervention classes that are broader in scope, the number of epidemiological trials to inform WHO recommendations was reduced. In doing so, however, the product diversity within a class increased considerably, raising the question as to whether products grouped within a specific class perform similarly to the first-in-class (FIC) product that established the intervention class and whether the WHO recommendation that was originally developed based on data for the FIC product continues to be applicable to the increasingly diverse group of products.

This uncertainty was recognized by WHO and its advisory groups as early as 2017 and, based on a technical consultation, WHO embarked on a process to explore the use of comparative efficacy to address these concerns *(4)*. A notice of intent to this effect was published by WHO in 2018 *(5)*, followed by a study protocol in 2019 *(6)*. The process was further explored by means of generating data for mosquito nets treated with a pyrethroid insecticide and the synergist piperonyl butoxide (PBO) *(7)*. For IRS, comparative efficacy data were used to expand the relevant WHO recommendation for IRS to neonicotinoid insecticides in 2017, and the need for comparative data is explicitly referenced in the associated preferred product characteristics *(8)*.

Based on these encouraging practical experiences and in the context of an ever-increasing diversity of vector control products, the WHO Malaria Policy Advisory Group recommended that WHO further advance the implementation of comparative efficacy assessments (using a method called "non-inferiority assessment") to expand on this exploratory area *(9,10)*. In 2023, the Malaria Policy Advisory Group reiterated its earlier guidance that comparative efficacy assessments of entomological data are required for all products other than the FIC products that generated the epidemiological data used to establish an intervention class *(11)*. In line with this guidance, the Global Malaria Programme launched a call for data and convened an expert group to assess the latest set of comparative efficacy data submitted to WHO and further engaged with the Guidelines Review Committee to discuss and evolve the utility of these data within the guidelines development process. Methodological recommendations from the 2021 *(7)* and 2023 *(12)* technical consultations have informed the present update to the data requirements and protocol published here.

1.1 How the comparative efficacy guidance fits with the WHO evaluation pathways and guideline development

For assessment of public health value, two well conducted, adequately powered trials with epidemiological end-points are needed. These data are used, sometimes along with other eligible studies in the public domain, to develop the overarching WHO recommendation for that intervention class. Once the intervention class has been established, other products within that class (called second-in-class (SIC) products) do not need to conduct epidemiological trials; however, they do need to demonstrate that they are non-inferior to the FIC intervention in terms of entomological efficacy.

Therefore, entomological non-inferiority assessments are conducted in place of epidemiological trials, with the overarching goal of providing reassurance to WHO with respect to the applicability of the recommendations, as stated in the *WHO guidelines for malaria (13)*, to new SIC products.

1.2 Intervention classes affected by this guidance

At the time of publication, comparative assessments of entomological data are required for all products other than the FIC products that generated the epidemiological evidence used to develop the WHO recommendation, thereby establishing that intervention class. The only exception is with pyrethroid-only insecticide-treated nets (ITNs), which are in the process of being replaced by more effective ITNs. All other current and future intervention classes, once established, will be required to follow this guidance, which will be modified to accommodate the specificities of new interventions currently undergoing epidemiological evaluation.

1.3 Rationale for comparative efficacy assessment using non-inferiority analysis

Entomological trials conducted in experimental huts measure proxies of likely epidemiological impact by measuring the effects of an intervention on end-points that are closely related to mosquito vectorial capacity *(14)*. Mosquito mortality is the end-point used for comparative efficacy assessments of all ITN and indoor residual spraying (IRS) interventions that use insecticides primarily designed to kill mosquitoes. For interventions designed to sterilize mosquitoes, additional reproductive end-points will need to be used. Comparative efficacy assessment focuses on the question: Compared to the FIC product for which there is evidence of public health benefit, is the candidate product not unacceptably worse with respect to the primary entomological end-point? The non-inferiority analysis evaluates whether the performance of the candidate product on the primary end-point is lower than that of the FIC product within the established margin, using standard WHO experimental hut methods and regression analysis *(6,7)*.

The design allows for meaningful comparison between the candidate and FIC products, using the same methods at the same site at the same time. This is important because absolute product performance will vary by location due to differences in local vector species and experimental hut design. To implement non-inferiority assessments of candidate SIC products, WHO has conducted work to find a balance between the non-inferiority margin needed to ensure that SIC vector control products are not unacceptably worse than the FIC product and the number of nights that is feasible for a hut trial, as part of studies conducted routinely to generate data required for prequalification.

1.4 Selection of the non-inferiority margin

The trial of a new vector control product should be designed to clearly demonstrate that the candidate product is not unacceptably worse (non-inferior) than the FIC product for which there is evidence, generated from randomized controlled trials (RCTs), of public health benefit in reducing malaria. This non-inferiority margin was defined based on mathematical modelling in which simulations were run to estimate how many additional malaria cases would occur if a product was between 5% and 30% inferior to the existing standard of care. It was decided that new products must be no more than 10% worse, and comparative efficacy trials need to be conducted at a sufficient level of quality to clearly demonstrate this. The maximum 10% absolute percentage difference was converted to a fixed odds ratio (OR) of 0.7 to enable comparative efficacy assessments from multiple trials in different geographies, where absolute values of product efficacy may vary using the same metric, i.e. the probability of a new product killing a mosquito is no more than 9% lower than that of an FIC product

that kills 50% of mosquitoes *(6)* (if 50% are killed by the FIC product and 41% are killed by the new product, the OR for mortality would be 0.7 [(41/59)/(50/50) = 0.7]). Following a review in 2023, this has been revised to use a variable OR that translates at each percentage mortality to a fixed percentage difference in mortality of 7% between the mortality induced by the FIC product and the lower bound of the 95% confidence interval (CI) of the mortality induced by the candidate product *(**Error! Reference source not found.**)*. To calculate the FIC mortality, the unadjusted value is estimated from the data, rather than through a regression model to select the relevant value of delta (non-inferiority margin).

1.5 Cost and conduct of comparative efficacy trials

It was demonstrated from an initial set of non-inferiority studies on pyrethroid-PBO nets *(7)* that non-inferiority assessments can provide a relatively easy and extremely cost-effective means of determining the entomological performance of products relative to an active comparator, using existing WHO methods *(15,16)*. Two trials were conducted – one in Côte d'Ivoire (122 days) and one in the United Republic of Tanzania (36 days) – with costs in line with those of the standard experimental hut trials that are routinely conducted as part of the data package required for WHO prequalification of new vector control products. Studies designed to evaluate new ITNs, IRS or other vector control interventions for WHO prequalification should be adequately powered to enable comparative assessment of the data in order to avoid the need for additional trials.

This protocol sets out the requisite procedures for a non-inferiority study in experimental huts and thus complements existing WHO testing guidelines for the evaluation of new vector control products.

1.6 Principles of non-inferiority determination in comparative efficacy trials

- Non-inferiority determination is a comparative assessment drawing on indirect (entomological) evidence to provide a certain level of reassurance of the likely public health benefit of all products within a product class other than the FIC product(s) that directly generated epidemiological data to demonstrate such impact.
- No additional studies beyond those needed to generate data for the WHO prequalification dossier should be required to generate the data to enable non-inferiority assessment.
- The results from the non-inferiority assessment cannot be used as a label claim.
- Non-inferiority assessments are intended to support the decision-making processes underpinning the sourcing of vector control products by WHO Member States and procurement agencies.
- Non-inferiority is not used as a measure of product quality; it is used only to assess the entomological efficacy of the product.

1.7 Aim of comparative efficacy assessment

The overarching aim of comparative efficacy assessment is to decide whether a candidate vector control product should be considered to be covered by an existing WHO policy recommendation for a vector control intervention class, using entomological data. This involves:

- conducting non-inferiority analysis to show that the entomological efficacy of a candidate vector control product is no worse than that of the FIC product by more than a pre-specified non-inferiority margin;

- where a claim of superiority to the standard of care is relevant, demonstrating that the candidate product and the FIC product are superior to the current standard of care (e.g. pyrethroid-only ITNs); and
- reporting the trial end-points and statistics in a standardized way.

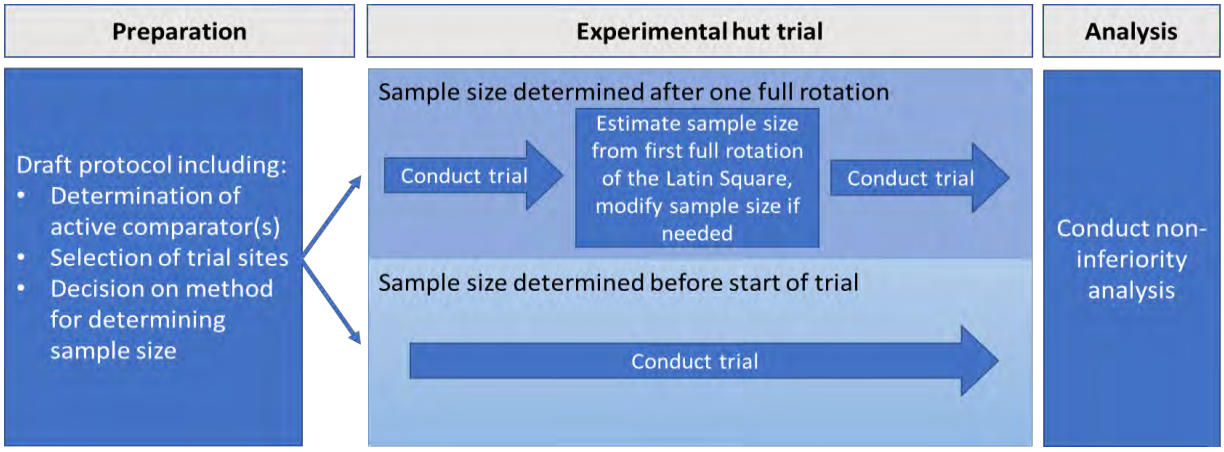
Any data submitted should have been generated in line with WHO guidance *(6)*, including the following:

- Data have been generated from at least two independent studies to inform the assessment.
- Data have ideally been generated in different geographical regions.
- Data from each study have been analysed separately.

2. Conduct of comparative efficacy assessments

Comparative efficacy evaluations are conducted as summarized in Fig. 1. An active comparator is selected (section 3.1) to which the candidate will be compared, and at least two trial sites are selected for the conduct of the evaluation (section 2.2). The trial is conducted using standard end-points (section 2.3) and adequate replication (section 2.4) with a standard study procedure (section 2.5) and appropriate treatment arms (section 2.6) to ensure that non-inferiority of the candidate to the active comparator may be reliably inferred using a standardized analysis (sections 3.1–3.4) and reporting (sections 3.5–3.6). The comparative efficacy trial is supported by quality checks of all vector control products to ensure that all trial arms are of a sufficient standard and to support inference of results (section 4).

Fig. 1. Outline of the process for comparative efficacy assessment



2.1 Selection of the active comparator

The active comparator for the non-inferiority trial should be the FIC product that demonstrated public health value by means of RCTs with epidemiological end-points. This is irrespective of the number of other products in the class. That is, the candidate product should ideally be compared to the FIC product and not to another SIC product in order to facilitate the comparison of products to a single standard with proven public health value and prevent "bio-creep" *(17)*. However, given the potential difficulty of obtaining FIC products for evaluation in non-inferiority studies, it may be acceptable for investigators to use an SIC product as the active comparator for comparative testing, if appropriate justification for this choice is provided.

Therefore, there are four options for active comparator products, listed below in order of preference:

- the FIC product of that intervention class;
- any SIC product for which epidemiological evidence of public health benefit is available from RCTs;
- a product that has shown superiority to the FIC product with respect to the primary end-point in entomological trials; and
- in the event that no SIC product has shown superiority to the FIC product and the FIC product is unavailable, the best performing product among the SIC products.

The selection of a specific active comparator needs to be justified in the study report. A table of suggested active comparators is presented in Annex 1.

If sufficient epidemiological evidence is available for two or more products in the same intervention class, the study investigator can choose between active comparators that have the requisite epidemiological data. The selection of products should be prioritized based on similar active ingredients (AIs), design, method of production and insecticide delivery method (e.g. incorporated vs coated for ITNs), if available.

Preferably, the active comparator should be obtained from the manufacturer to ensure that it has not degraded through poor storage, which could result in inconsistent performance. A certificate of analysis of any active comparator should be supplied with the report to assure the quality of the product used in the comparative efficacy study.

In addition to the active comparator, all studies are required to include a negative control arm (i.e. an untreated net for ITNs or water spray for IRS) to i) verify that the conduct of the experimental hut trial is of sufficient quality, and ii) estimate the natural mortality during mosquito holding or, where relevant, changes in blood feeding that are induced by the test product. Trials in which the overall mortality in the control arm (over the duration of the study) is > 10% at 24-hour holding or > 20% for longer holding times need to be investigated and the trial repeated.

2.2 Site selection

Suitable sites for undertaking the study need to meet multiple criteria. First, there should be enough huts of the same design to enable all arms of the experiment to be run simultaneously. Quality-assured (Good Laboratory Practice (GLP)-compliant) test facilities are required for vector control product evaluation. Each site also requires sufficient entomological data from a recent study or a pilot study to ensure accurate power calculations. Such data include recent mosquito densities, mortality induced by the FIC product, and sources of variability, e.g. hut, volunteer and nightly heterogeneity *(18)*. If these data are unavailable, then the study power is checked with an interim power estimation during the trial (e.g. after one full rotation of the treatments or 49 nights for a 7x7 trial). The study protocol should be modified if additional rotations are required to achieve power.

Mosquito species composition and the resistance status of each major vector species at the study site must be characterized. This should include determination of the phenotypic resistance frequency and molecular characterization of the resistance mechanisms, especially metabolic mechanisms for products designed to counteract this type of resistance mechanism. For instance, if the candidate is a pyrethroid-PBO ITN, the local mosquito population should be assessed for P450 expression prior to exposure to PBO, ensuring that assays are conducted under optimal temperature with the correct holding time *(19)*. For insecticides that show < 90% induced mortality in a discriminating dose bioassay, the intensity of resistance should be quantified using the procedures outlined in existing WHO guidance *(19)*.

If *Anopheles* vector species complexes are present, data should be collected and analysed to determine the dominant species. However, a pooled analysis of species complexes may also be performed, if justified, ensuring that the insecticide resistance status of each subspecies is evaluated and reported.

Two independently powered studies should be conducted in two sites with differing vector populations and/or resistance status, such as one in West or Central Africa and one in East Africa. Study protocols should be registered with WHO prior to the start of studies.

If the results of the two initial studies are inconsistent in terms of whether the impact on the primary end-point(s) demonstrates non-inferiority, a third replicate trial will be required. Studies in which the candidate product fails to show benefit over the negative control/standard comparator (depending on the claim of the product) will need to be repeated at additional sites. No more than three studies should be conducted, and non-inferiority needs to be demonstrated in two distinct locations: Products fail to demonstrate non-inferiority if they show inferior or inconclusive results in two of the three comparative efficacy studies. These products will be required to undergo further product development to enhance their performance, or they will need to provide epidemiological evidence for assessment of their public health value.

2.3 Study design

2.3.1 Primary end-points

In comparative efficacy studies, a new product must show non-inferiority to the active comparator with respect to the primary end-point. The measurement and assessment of the primary end-point should be informed by the entomological mode of action of the intervention class (Table 1). For the current intervention classes, the end-point selected has been informed by the RCTs in which the FIC products were assessed, as this end-point has the greatest impact on epidemiological outcomes *(14)*. Absolute values should be given for the primary end-point (i.e. not corrected for the untreated control).

Mortality

When the primary end-point is mortality, the study should record the 24-hour mosquito mortality for the duration of the trial, unless there is a priori justification for using longer holding times up to 168 hours (e.g. based on the mode of action of the candidate product), provided that control mortality is sufficiently low (< 10% at 24 hours and < 20% at longer holding times) to warrant the inclusion of the data. Mortality at each holding time should be presented in the report, and non-inferiority analysis should be performed on the primary mortality end-point (Table 1). Justification for the selected holding time based on the mode of action of the FIC and SIC products should be included in the report, and all study arms should be measured at the same holding time(s).

Fecundity

When the primary end-point is a reduction of fecundity, the study should record the proportion of mosquitoes that were alive with viable eggs at a set time (e.g. 72 hours) after the collection of blood-fed females from an experimental hut. Fecundity may also be measured by counting the number of eggs laid by each blood-fed mosquito that remained alive long enough to complete egg development and lay eggs, or by dissecting mosquitoes to look for viable eggs. Data must be checked against an untreated control to ensure that it is the active component of the test product that is responsible for reducing mosquito fecundity, and not other factors.

Table 1. Non-inferiority end-points and when to measure them according to the chemical mode of action (for ITNs and IRS only)

Chemical class	Chemical mode of action	Example chemistry	Note	Primary end-point	Holding time before measurement	Additional end-points	Holding time before measurement	Additional end-points	Holding time before measurement	Quality assurance bioassay
Sodium channel modulators (pyrethroids)	Nerve action	Pyrethroids Pyrethroids with synergist PBO	Not applicable for pyrethroid-only ITNs	Proportion dead	24 hours			Proportion blood-fed	24 hours	Cone test
Acetylcholinesterase (AChE) inhibitors (carbamates, organophosphates)	Nerve action	Pirimiphos-methyl Bendiocarb		Proportion dead	24 hours			Proportion blood-fed	24 hours	Cone test
Nicotinic acetylcholine receptor competitive modulators (neonicotinoids)	Nerve action	Clothianidin	AI is slow-acting so longer holding times are needed	Proportion dead	72 hours		24 and 48 hours Longer holding times in intervals of 24 hours up to 168 hours	Proportion blood-fed	24 hours	Cone test
Uncouplers of oxidative phosphorylation via disruption of the proton gradient (pyrroles)	Energy metabolism	Chlorfenapyr	Requires insects to be metabolically active during testing and AI is slow-acting so longer holding times are needed	Proportion dead	72 hours		24 and 48 hours Longer holding times in intervals of 24 hours up to 168 hours may be included	Proportion blood-fed	24 hours	Tunnel test
GABA-gated chloride channel allosteric modulators (meta-diamides and isoxazolines)	Nerve action	Broflanilide	AI is slow-acting so longer holding times are needed	Proportion dead	72 hours		24 and 48 hours Longer holding times in intervals of 24 hours up to 168 hours may be included	Proportion blood-fed	24 hours	Cone test
Juvenile hormone mimics (pyriproxyfen)	Growth regulation	Pyriproxyfen	Inhibits the development of viable eggs/larvae	Proportion with viable eggs or number of eggs per female	72 hours	Proportion dead	24, 48 and 72 hours	Proportion blood-fed	24 hours	Cone test

2.3.2 Secondary end-points

The following end-points should be reported in all instances and measured over the same duration as the primary end-point for all trial arms (see glossary for definitions):

- proportion of blood-fed mosquitoes; and
- proportion dead at 24-hour intervals up to the longest holding time used in the trial.

2.4 Sample size considerations

Non-inferiority analysis requires estimates of the OR to be as precise as possible to enable clear classification and avoid, to the extent possible, requests for additional studies. Sample size is estimated as the number of replicates required to precisely measure the point estimate of the primary end-point within a predefined non-inferiority margin. For WHO assessments of non-inferiority, a margin of 7% of the absolute difference in measures of the primary end-point between treatment arms should be used. Justification for the sample size should be presented in reports alongside study results. The requisite sample size will depend on the number of replicates, the absolute entomological efficacy of the intervention tested and other variability inherent in the study.

The study power measures the chance that the trial will demonstrate the non-inferiority of the new product to the active comparator if the true efficacy of the new product is the same as that of the active comparator. Study power can be estimated by simulating candidate and active comparator products that have the same underlying efficacy and then determining the percentage of runs that correctly classify the candidate as non-inferior. The study should be designed to have a power of 80% (i.e. $\beta = 0.2$).

Sample size calculations can be estimated through simulation. The primary analysis makes the assumption that the average impact of the intervention over its lifetime can be used as a single end-point, thus simplifying power calculations. For experimental hut trials, variability includes the number of mosquitoes collected per hut per day, as well as differences between huts, between sleepers and between observations. These factors will vary by setting and over time, so it is important to use data from recent hut trials or pilot studies to parameterize the sample size (power) calculations. If recent data are unavailable to parameterize the power calculations, the number of replicates can be estimated after two rotations of the Latin square (LS) (e.g. for a 7x7 LS, this would be after 14 experimental nights). If the sample size is too small and the study needs to be run for longer, this should be documented as an amendment to the protocol. It is also recommended that the power of the study be checked after two rotations in case additional replication is needed to determine non-inferiority *(7)*, e.g. if the number of mosquitoes collected per night or the impact of the active comparator on the primary end-point is lower than expected.

The steps to conduct sample size calculations for comparative efficacy assessments using the example of mortality are as follows:

1. Estimate mosquito mortality observed for the active comparator.
2. Estimate the average number of mosquitoes collected per hut and the variability in mosquito density per day.
3. Estimate the variability in hut, sleeper and daily observations.
4. Define the ideal number of huts to use, days per week the experiment is to be run (e.g. seven days for a 7x7 design, followed by a break before the next treatment is allocated in the case of ITN trials) and the number of LS rotations for the experimental hut trial (e.g. for a 7x7 design, a minimum of one full rotations, 49 nights). These values should represent the

minimum stipulated above (and in the WHO guidance on experimental hut studies *(15,16)*), but additional investigational arms (i.e. more than one hut per arm) or rotations may be required to ensure sufficient power for the non-inferiority analysis.

5. Use the data from steps 1–4, together with the defined non-inferiority margin based on a 7% absolute difference in mosquito mortality between the active comparator and the candidate product, to simulate theoretical experimental hut trial results for all trial arms (assuming that the percentage mortality follows a binomial distribution). To estimate study power, the true mortality of the test product (i.e. the underlying actual probability that a mosquito will die) should be the same as that of the FIC product, i.e. the candidate product is truly no worse than the active comparator.
6. Fit the logistic regression model outlined in section 3.1 to simulated data and determine whether non-inferiority has been shown.
7. Repeat steps 5–6 1000 times and calculate the percentage of times non-inferiority is demonstrated. Record this as study power.
8. Repeat steps 5–7, adjusting the number of replicates used (i.e. increasing the number of huts for each trial arm or the number of rotations) until the desired power of > 80% has been achieved.

2.5 Conduct of experimental hut trials

2.5.1 Experimental huts

Experimental huts enable evaluation of a range of vector control products under controlled conditions that resemble how mosquitoes would enter a human habitation and interact with the product under normal use. Experimental huts have structural features that enable the collection of mosquitoes that have entered the huts, and it is possible to measure multiple end-points. There are several kinds of experimental huts in use, and all have common design features:

- Within a trial site, all huts are identical and located in a single location for each evaluation.
- Huts have entry points (eave gaps or entry slits) that allow host-seeking mosquitoes to enter and search for blood hosts, but that minimize egress so that mosquitoes are retained.
- Huts are positioned in proximity to mosquito breeding sites to allow for a uniform rate of mosquito entry into each hut.
- A water-filled channel surrounds each hut to prevent entry of ants that would scavenge for incapacitated or dead mosquitoes, which would result in underestimation of mosquito mortality.
- Each hut has traps at exits (eaves, windows or verandas) to capture exiting mosquitoes, enabling estimates of the exit rates of mosquitoes.
- Huts are generally small in size (< 3 m x 3 m) to simplify mosquito collection.

2.5.2 Mosquito scoring and handling

Each morning of the study, live and dead mosquitoes are collected using standard procedures (e.g. aspiration of mosquitoes from inside the net, on the floor, on the walls and in exit traps or verandas) and recorded on a standard format data sheet (Annex 2). Mosquitoes from each hut are then sorted morphologically by species and location (inside the hut, in exit traps or verandas) and classified as dead unfed, dead fed, alive unfed and alive fed. Where cryptic species occur in the area, appropriate molecular methods should be used for species identification at the end of the holding period.

For delayed mortality assessments, live mosquitoes are placed in paper cups with no more than one mosquito per 20 cm^3 to prevent damage during close confinement. Ideally, holding conditions should

be 27°C ± 2°C with 75% ± 10% relative humidity and light–dark cycles appropriate for the species. The actual holding conditions should be recorded and reported. Mosquitoes are given access to a 10% sugar solution throughout the holding period. Delayed mortality is recorded after every 24 hours for up to 168 hours, depending on the mode of action of the product tested (Table 1).

If reproductive outcomes are to be evaluated, reduction of fecundity is the primary end-point for all ITNs that affect reproduction, while mortality is the secondary end-point. Reduction of fecundity is defined as the decrease in the proportion of females with viable eggs or decrease in the number of eggs produced by a surviving blood-fed adult female mosquito.

Fecundity is calculated by dissecting mosquitoes to look for viable eggs at a set time, usually 72 hours after the collection of blood-fed females from an experimental hut. It may also be measured by counting the number of eggs laid by each blood-fed mosquito that remained alive long enough to complete egg development and lay eggs. Data are checked against an untreated control to ensure that it is the product's AI that is inducing the change in mosquito fecundity.

2.5.3 Measurements to reduce heterogeneity and bias

Participant and investigator bias is reduced if interventions are identifiable by numeric code and unblinding occurs after data are entered and locked. Investigator bias is also reduced through the use of a predefined statistical analysis plan describing the analysis, which should be included as an annex to the study protocol.

Heterogeneity in nightly mosquito entry into huts is caused by the location of huts relative to breeding sites and sleepers' individual attractiveness to mosquitoes. The LS design compensates for these differences by ensuring that each sleeper is assigned to each treatment and hut the same number of times over the course of the study. Thorough training of mosquito collectors and those scoring mosquitoes is required to minimize operator bias.

2.5.4 Ethical and safety considerations

Where human participants are sleeping in the huts, institutional ethical approval for the study must be sought from the local ethical review board. Written informed consent will need to be obtained from each volunteer sleeper prior to their participation in the study, and possible adverse events should be monitored. In areas of malaria transmission, medically supervised studies in which sleepers are provided with malaria chemoprevention will minimize risk of infection during the trial *(20)*. Adequate training, product-specific information sheets and personal protective equipment should also be provided to anyone applying or handling pesticides.

2.6 ITN evaluation

Candidate ITN products must demonstrate non-inferiority to be considered covered by an existing WHO policy recommendation for a specific intervention class. Currently, ITNs must demonstrate non-inferiority up to 20 washes using the combined data for unwashed and washed (20 times) ITNs with respect to the mortality end-point. The data for the candidate product are compared to the combined unwashed and washed (20 times) data for the active comparator product in the same experimental hut study. Additional candidate products can be added to the study design to evaluate more than one candidate product at the same time. A negative control arm (untreated bednet) is included to assure the quality of the study by monitoring mortality in the control and to calculate blood-feeding inhibition if required.

For products designed to counteract pyrethroid resistance, such as pyrethroid-PBO ITNs, the inclusion of a pyrethroid-only ITN in the trial is required to distinguish an additionally effective vector control product from one that may no longer be fully effective. This net will be the standard comparator; it

should be a WHO-prequalified product and should be selected based on the standard of care in the geographical region where the experimental hut trial is being conducted or where the product may be used. Currently, the standard comparator for trials of new ITNs is a pyrethroid-only ITN, given that this continues to be the most widely used vector control intervention for malaria.

The following study arms are currently recommended for products to evaluate a candidate dual AI net in an area of pyrethroid resistance:

1. Candidate ITN unwashed (candidate product)
2. Candidate ITN washed 20 times (candidate product)
3. Active comparator unwashed (see options provided in section 2.1)
4. Active comparator washed 20 times (see options provided in section 2.1)
5. Pyrethroid-only ITN unwashed (standard comparator)
6. Pyrethroid-only ITN washed 20 times (standard comparator)
7. Untreated net unwashed (negative control).

2.6.1 Experimental hut procedure for ITNs

To ensure standardization of experimental hut trials across the numerous trial sites and ITNs being tested, all ITNs tested must be evaluated before and after a standard number of washes, as outlined in WHO guidelines *(15)*. Nets should be tested both in an unwashed state and after being washed 20 times.

The interval between each wash is also important and needs to be standardized for the duration of the washing process for each ITN. The wash interval is based on the time it takes for the bioavailability of the AIs and synergists to be restored on the surface of the net after washing (regeneration time). As the regeneration time will depend on the chemistry of the ITN, it should be determined according to established procedures *(15),* using an appropriate mosquito strain (e.g. for PBO ITNs, a resistant strain with a monooxygenase-based resistance mechanism). If an ITN is WHO-prequalified, the published wash interval must be used to ensure comparability between trials. The washing process may take several months and should be coordinated such that all the nets being used in the trial undergo their 20th wash synchronously and are ready to start the trial at the same time. It should be noted that the experimental hut trial should start only after the nets have been allowed to regenerate for two weeks following the final wash.

Before the field study commences, all nets are deliberately holed following standard guidelines *(15)*. Six holes (4 cm x 4 cm) are cut in each net: two holes in each of the long side panels, and one hole at each end (head and foot end). Holes are made halfway up the side of the ITN, which can be most easily measured as 75 cm from the top seam.

The procedure for comparative efficacy evaluation of ITNs is identical to that outlined in WHO guidance *(15)*. Treatments are initially randomly allocated to the experimental huts, and then sleepers and treatments are rotated through the experimental huts, using a partially or fully randomized LS design. Ideally, by the end of the trial, each treatment will have been tested in each experimental hut with each sleeper the same number of times. The exact study design will depend on the number of study arms. Illustration of the method is provided in existing WHO guidance *(15)*. To increase the precision of estimates through improved sample sizes, more than one LS can be conducted at a time at the same site, and data combined for analysis. For example, two 7x7 LSs can be conducted in 14 huts at the same time over 49 nights of data collection.

Individuals will rotate through huts on a nightly basis, and the treatment will rotate after each sleeper has spent one night with each treatment (e.g. in a 7x7 LS, the treatment will move after seven nights in one hut). When rotating treatments, a window of at least 24 hours is required so that the huts can

be thoroughly cleaned and aired out. This will minimize any carry-over effects between treatments. A minimum of six replicate ITNs per arm should be used to account for between-net heterogeneity. If more than six nights are required for the LS, then the individual ITNs can be rotated, ensuring that each ITN is tested a similar number of times.

2.6.2 IRS evaluation

Candidate IRS products must demonstrate non-inferiority to be considered covered by an existing WHO policy recommendation for a specific intervention class. Currently, IRS products must demonstrate non-inferiority to the active comparator, using combined data for all substrates tested in the trial on the mortality end-point at the longest duration of efficacy. At minimum, the non-inferiority of IRS should be assessed at three months, as this is the minimum duration of efficacy required for products in this intervention class. For products with longer residual efficacy, trials should be conducted until 80% efficacy of one of the two products being compared is no longer achieved or at a justifiable predefined residual efficacy, e.g. determined by laboratory assays. The minimum estimated duration of efficacy should be considered in calculating the sample size and estimating the number of replicates needed for the trial.

IRS products must demonstrate non-inferiority to a FIC product, which may include IRS products of different insecticide classes, i.e. with different chemical modes of action, provided that mortality is measured for all products at the same holding time. For IRS, a standard comparator may be used, although it is not a requirement. For IRS, the standard comparator should be a WHO-prequalified product (but not a pyrethroid insecticide); it should be selected based on the standard of care in the geographical region where the experimental hut trial is being conducted or where the product may be used. Data must be reported for a 24-hour holding time, but longer holding times can be used for products with slower modes of action, if required, provided that the control mortality is acceptable (< 20%). The following study arms are required at minimum:

1. Water (negative control)
2. Active comparator with epidemiological evidence of impact from RCTs, with a similar expected duration of residual efficacy (see options provided in section 2.1)
3. Candidate IRS (candidate product).

In this scenario, additional arms may also be included to incorporate insecticide classes currently used or planned to be used for IRS in the country. The material used for the walls of the structures to be sprayed (e.g. mud, concrete, wood, etc.) will affect the performance of the product. Therefore, the selection of the substrate should be justified based on the common housing materials in the region where the product is to be used.

Because IRS treatments cannot be rotated between huts, the use of four huts per treatment arm is recommended to overcome the spatial heterogeneity between huts. Data quality is improved by increasing the number of huts per arm, and this should be considered in the power analysis at the start of the study. To this end, study arms, and the number of replicates per arm, should be maximized within the limits of practicality, depending on hut availability for the selected study site. If the number of huts in an experimental hut trial site is a limiting factor, the number of negative control huts can be reduced to one per substrate type.

2.6.3 Experimental hut procedure for IRS

IRS is applied to the walls of the hut in accordance with the manufacturer's usage instructions. The ceiling and doors are left unsprayed to avoid confounding the efficacy of the IRS on different substrates, e.g. mud walls and thatch roof. The ceiling and doors of the huts in IRS trials should be

covered with a material that reduces mosquito resting (e.g. stretched plastic) so as to maximize the likelihood of mosquitoes resting on the treated surfaces.

The quality of spraying in hut trials is an essential prerequisite for any comparative efficacy evaluation or study for WHO prequalification. Spray application must be within ±50% of the label-recommended target dose for the IRS product *(21)*, as determined through filter paper analysis (described below). Optimal spraying is achieved by employing well trained personnel *(22),* using calibrated compression sprayers with control flow valves and carefully calculating the concentration of insecticide in the tank prior to spraying. Gravimetric verification of the spray dose is recommended following i) the calculation of the hut surface area sprayed; ii) the weighing of spray tanks before and after spraying; and iii) the estimation of grams of solution applied per metre of surface in the huts. Chemical verification of the target dose is outlined in section 4. Treatments should be allocated randomly to the experimental huts. Sleepers should enter and leave the huts at predefined times each night and in the morning. Alternatively, yet less preferably, cows may be used as bait animals, and these animals should also be rotated each night. The exact design will depend on the number of study arms. Illustration of the method is provided in existing WHO guidance *(16)*.

3. Data analysis and management

3.1 Non-inferiority margin

The non-inferiority margin has been defined as a 7% difference in absolute efficacy between the FIC product and the candidate product. This margin is translated into an OR based on the absolute value of the primary end-point measured for the FIC product in each trial (Table 2).

Table 2. The non-inferiority margin based on a fixed difference of 7% between the candidate and the active comparator (FIC) products

Active comparator %	Lower bound of candidate CI if non-inferiority margin is 7%	Corresponding OR for a 7% non-inferiority margin (mortality)	Inverse of the OR corresponding to a 7% non-inferiority margin (fertility, blood feeding)
95	88	0.39	2.56
90	83	0.54	1.85
80	73	0.68	1.47
70	63	0.73	1.37
60	53	0.75	1.33
50	43	0.75	1.33
40	33	0.74	1.35
30	23	0.70	1.43

A variable OR that is calculated based on a fixed absolute difference of 7% has been adopted ***(Error! Reference source not found.)***. The fixed OR of 0.7 was introduced by WHO in 2019 *(6)* as a compromise between the risk of accepting an inferior product and the feasibility of conducting the trials. However, at low and high FIC mortality values (i.e. below 30% and above 80%), the fixed OR of 0.7 equates to a very small absolute difference in mortality, which could prevent products from reaching the non-inferiority margin, even if they may be highly efficacious in absolute terms. Consequently, after reviewing the data sets from a number of trials, the fixed absolute difference was deemed more justifiable. The mortality for the active comparator used to determine the non-inferiority margin is simply calculated as the unadjusted arithmetic mean proportion for the primary end-point measured from the complete data set for the active comparator arm (without differentiating by wash status or substrate). For non-binary end-points, i.e. number of eggs laid, the arithmetic mean proportion of

reduction in eggs laid (relative to the negative control) may be used in the same way to determine the non-inferiority margin.

Since higher mortality indicates a better product, a candidate product will be determined to be non-inferior in terms of mosquito mortality if the lower bound of the 95% CI estimate is greater than the non-inferiority margin (Table 2).

When the primary end-point is the proportion of mosquitoes that are fertile (or that are blood-fed), better products should have lower values. In this case, the OR is set at the inverse of the non-inferiority margin, i.e. (1/odds of non-inferiority margin). A candidate SIC product will show evidence of non-inferiority if the upper bound of the 95% CI estimate is lower than the non-inferiority margin (Table 2).

To ensure standardization of analytical approaches, a specific model must be used when performing the analysis and presenting the results for assessment of non-inferiority. To link the outcome variables to the intervention and covariates, generalized linear regression models should be used. The choice of model will depend on the end-point(s) under investigation. For binary end-points, such as the proportion of mosquitoes dying or the proportion that are fertile, a logistic model is appropriate. For outcomes that are counts, such as the number of eggs laid, a Poisson or negative binomial model may be more appropriate.

All covariates should be categorical fixed effects and the active comparator should be used as the reference intervention (intercept). For experimental hut trials, covariates include treatment, huts (if multiple huts are used per treatment arm), sleepers and night as fixed effects, because these factors are sources of systematic variability that are accounted for in the experimental design.

3.2 Primary end-point

To generate a single estimate of efficacy for the primary analysis, data for both washed and unwashed nets of a single product should be analysed together to give an estimate of overall product performance over its lifetime in the field. Combining the two arms increases replication in the analysis and consequently the precision of the estimates. Similarly, for IRS, the analysis should be conducted over the full duration of the expected product efficacy (i.e. residual efficacy).

Three analyses are conducted for ITNs:

- primary analysis of the primary end-point: all data (unwashed and washed for ITNs); wash status is included as a fixed covariate in the primary analysis;
- secondary analysis of the primary end-point for unwashed ITNs; and
- secondary analysis of the primary end-point for washed (20 times) ITNs.

Three analyses are conducted for IRS:

- primary analysis of the primary end-point: all substrates at the longest duration of efficacy; substrate is included as a fixed covariate in the primary analysis;
- secondary analysis of the primary end-point for a single IRS substrate, e.g. mud, at the longest duration of efficacy; and
- secondary analysis of the primary end-point for a single IRS substrate, e.g. cement, at the longest duration of efficacy.

For IRS, if the duration of efficacy exceeds three or six months, then the non-inferiority for all substrates combined and each individual substrate should be determined at three months and at six months to understand the decay of the IRS product over its residual life.

The estimated effect of the intervention, e.g. % mortality and 95% CIs for each arm, and the combined results for washed and unwashed nets should be reported in addition to the OR or rate ratio (RR) and its 95% CI from the regression model (section 3.6). Results of all analyses may be presented together in forest plots (e.g. Fig. 2) and subdivided by species where relevant. An example of an analysis script is available in Annex 3.

3.3 Secondary end-point

Secondary end-points, namely blood feeding for ITNs, mortality at holding times other than the holding time used for the primary end-point or mortality for ITNs that have fertility as the primary end-point, should also be analysed and presented as per section 3.6. This information will enable greater understanding of the full mode of action of a product but will not be used for decision-making.

3.4 Non-inferiority test

A candidate product must show non-inferiority to the active comparator product with respect to the primary end-point to become part of the intervention class covered by a WHO policy recommendation.

The candidate product is deemed non-inferior if the following criteria are met:

- The primary end-point is mosquito mortality (e.g. for ITNs and IRS) and the lower 95% CI estimate of the OR between the candidate product and active comparator product is greater than the OR corresponding to a 7% non-inferiority margin (see Table 2).

or

- The primary end-point is mosquito fertility (e.g. for sterilizing ITNs) and the upper 95% CI estimate of the OR between the candidate product and active comparator product is less than the inverse of the OR corresponding to a 7% non-inferiority margin (see Table 2).

and

- The candidate product is classified as superior to the negative control or standard comparator in terms of mosquito mortality if a significantly higher proportion of mosquitoes have died at the 5% significance level (i.e. $P < 0.05$). The mortality end-point is used for evaluation of both ITNs and IRS. The choice of whether the candidate product should be compared to a control or to the current standard of care will depend on the product and should be justified (Annex 1).

3.5 Data reporting

Primary and additional end-points should be reported for each trial arm with appropriate measures of centrality and dispersion, e.g. arithmetic mean % mortality with 95% CI. In addition, the total numbers of female mosquitoes collected, total dead and total blood-fed should be reported, as well as the number of hut nights per trial arm (Annex 4). All raw data should be provided upon request to WHO using a standard format (Annex 2), recording the number of mosquitoes caught that are alive unfed, alive fed, dead unfed and dead fed per hut night in each arm per holding time.[1]

[1] Data access is restricted to WHO staff only prior to publication of the primary trial data without explicit permission of the primary investigator.

IRS studies should report monthly residual efficacy in free-flying mosquito populations and with confirmatory cone bioassays (for neurotoxic insecticides) to demonstrate the number of months that residual efficacy exceeded 80%.

3.6 Data presentation

Data should be presented in a standard way to ensure its comparability and easy interpretation. Data should be presented in both tabular and graph form. Examples are given below.

For the format of data presentation in tables, fictional examples are given in Table 3 for ITNs and Table 4 for IRS.

Table 3. Point estimates of pooled data from unwashed and washed ITNs

Outcome	Product	Role in study	Mean %	95% CI
Primary: Mortality (72 hours)	Untreated net	Negative control	1.2	0.2–2.1
	Pyrethroid-only ITN	Positive control	17.3	15.3–19.3
	PBO FIC	Active comparator	79.0	76.8–81.2
	PBO candidate net	Candidate	75.8	73.4–78.2
Secondary: Blood feeding	Untreated net	Negative control	80.2	76.4–85.3
	Pyrethroid-only ITN	Positive control	50.6	48.0–53.2
	PBO FIC	Active comparator	26.2	23.8–28.6
	PBO candidate net	Candidate	34.5	31.9–37.1

Table 4. Point estimates of 72-hour mosquito mortality outcomes for the respective products tested against IRS products

Outcome	Product	Role in study	Substrate	Point estimate	95% CI
Primary: Mortality (72 hours)	FIC IRS	Active comparator	Overall	44.2	40.7–48.1
	Candidate IRS	Candidate	Overall	57.5	52.9–61.0
	Control	Negative control	Overall	2.1	1.2–3.0
	FIC IRS	Active comparator	Concrete	43.8	39.1–48.5
	Candidate IRS	Candidate	Concrete	57.0	54.2–59.8
	Control	Negative control	Concrete	1.6	1.2–2.0
	FIC IRS	Active comparator	Mud	44.6	33.9–49.2
	Candidate IRS	Candidate	Mud	57.9	55.0–60.7
	Control	Negative control	Mud	2.6	2.2–3.0

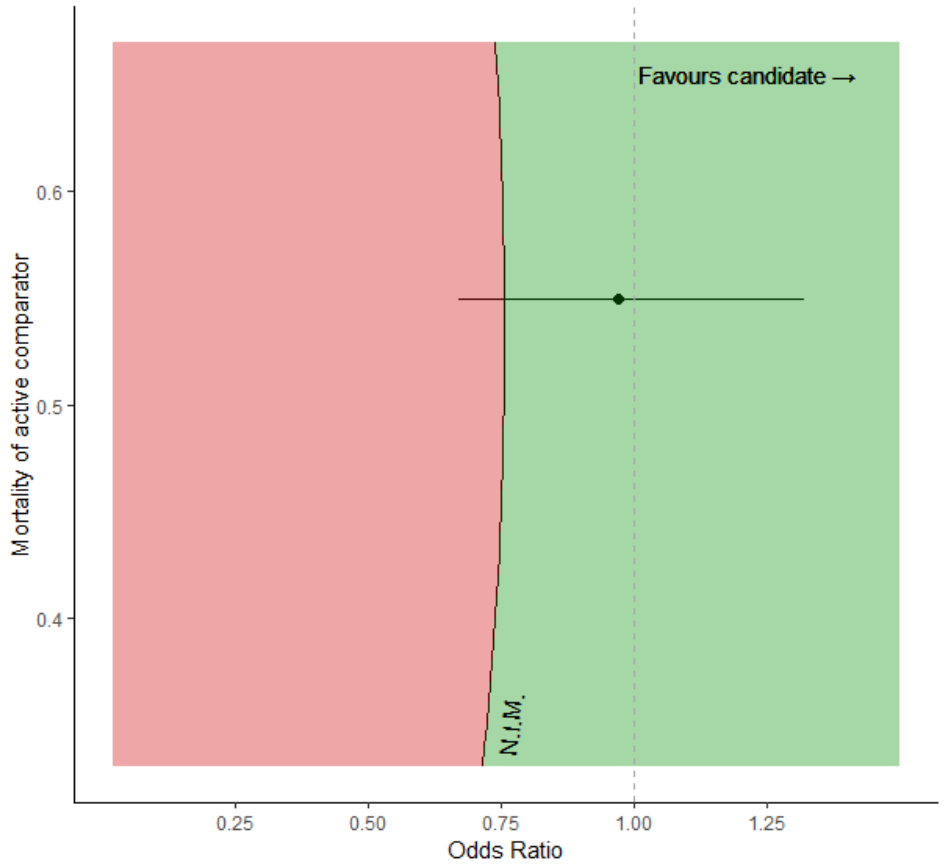
To present the outcomes of non-inferiority evaluations, the data can be presented as in Table 5.

Table 5. Fictional presentation of ORs for a candidate IRS compared to an active comparator IRS to assess non-inferiority

Outcome	Active comparator	Candidate	Substrate	OR	95% CI	Target outcome	Test outcome
Primary: Mortality (72 hours)	FIC IRS	Candidate IRS	Overall	1.08	0.97–1.22	Non-inferior	Non-inferior
Primary: Mortality (72 hours)	FIC IRS	Candidate IRS	Concrete	1.16	1.01–1.35	Non-inferior	Non-inferior
Primary: Mortality (72 hours)	FIC IRS	Candidate IRS	Mud	1.01	0.87–1.16	Non-inferior	Non-inferior

Finally, when presenting the ORs and non-inferiority margins graphically, the studies can be plotted on a graph with the OR on the x-axis and the mortality in the active comparator arm on the y-axis. An example is provided in Fig. 2.

Fig. 2. A fictional example of a graphical presentation of the ORs of mortality from an experimental hut trial of IRS conducted using the new fixed difference non-inferiority margin



Note: As the non-inferiority margin corresponds to a fixed difference of 7%, its value varies when expressed as an OR.

4. Quality assurance bioassays

Additional bioassays may be conducted for quality control to confirm the continued efficacy of a product under evaluation, but data from this testing are not used as part of the comparative efficacy assessment. Confirmatory forced contact bioassays are conducted as part of the experimental hut testing of ITNs and IRS to assure the quality of the vector control products used and their correct application *(15,16)*. For products designed to combat vectors with metabolic resistance, exposure at the time in the circadian rhythm when mosquitoes are metabolically active will give the most representative estimates.

Products containing AIs that induce rapid knockdown and mortality or AIs that affect reproduction *(23)* should be evaluated using cone tests. Products that contain pro-insecticides that are metabolized into their active form by mosquito detoxifying enzymes are more suited to evaluation using tunnel tests or other suitable laboratory assays in which mosquitoes are metabolically active e.g., free-flying (Table 1).

For ITNs, multiple steps are necessary, and all testing should be performed on supplemental nets that are from the same production batches as those in the trial and have been treated the same way. Normally, three production batches are used. Bioefficacy tests for ITNs involve checking ITN quality upon receipt, and efficacy testing post-washing and post-hut testing. First, three unwashed ITNs from each treatment arm (one from each batch) are checked for bioefficacy before any experimental hut testing is performed. Next, it is necessary to check that washing has been performed correctly; therefore, bioassays are performed on two washed nets from each washed arm that have completed the required 20 washes in accordance with the protocol, alongside those that have been prepared for the experimental hut testing. Finally, bioefficacy tests are performed again on two of the actual nets used in the trial, one for each treatment arm.

For IRS, forced contact bioefficacy testing (cone tests) is conducted on sprayed surfaces each month following IRS application. These tests should be continued on a monthly basis until the mortality of the mosquitoes in contact with the treated surface drops below 80% in order to determine the longest duration of efficacy. Some insecticides (e.g. uncouplers of oxidative phosphorylation via disruption of the proton gradient, including chlorfenapyr) cannot be monitored using cone tests and therefore bioassays using free-flying mosquitoes are more appropriate *(24)*. Therefore, only the mortality of free-flying mosquitoes in experimental huts can be used to estimate the residual efficacy of these products.

In addition, the verification of insecticide target concentration may be conducted using pesticide residue analysis *(25)*. Samples should be analysed in a quality-assured chemical testing laboratory.

Bioassays should be performed during the study according to standard WHO procedures *(16)*. For each of the sampled nets (i.e., all trial arms sampled unwashed, washed and naturally aged in the field), five 25 cm x 25 cm pieces will need to be cut from positions 1 to 5 as per WHO guidelines *(16)*. ITN samples cut directly adjacent to each test piece are placed individually in aluminium foil, labelled with the appropriate information (ITN number, date, location of sample) and stored in a refrigerator at 4°C until its chemical analysis.

IRS verification is conducted using four 10 cm x 10 cm papers (Whatman® No. 1) that are attached (backed with foil and using pins to hold them slightly away from the wall to avoid run-off) at three different wall heights (top, middle and lower part of the wall), plus one randomly assigned height, and then removed after the spray activity. Spray quality is assessed in each experimental hut. After spraying, filter papers are placed individually in aluminium foil, labelled with the appropriate

information (hut number, substrate, date, wall location – top, middle or bottom) and stored in a refrigerator at 4°C until its chemical analysis. The spots on the walls where the filter papers were placed should be marked to avoid placement of cones on those untreated areas during subsequent cone bioassay tests.

5. Registration of trials

To avoid the cherry-picking of positive trial results, all trials should be registered in a WHO registry of non-inferiority studies prior to undertaking the trials. Tests can only be done at GLP-accredited sites or, in the interim, performed according to GLP standards at sites undergoing certification. Sites are responsible for registering studies in their own registry that can be accessed by WHO upon request, with the aim of avoiding the failure to report negative trial outcomes. Other mechanisms to ensure such transparency should be explored, analogous to the requirement to register clinical trials. Trial sites selected to undertake the studies should be appropriate for the question being asked. Site selection should be justified based on a baseline assessment of class-relevant parameters (e.g. P450 resistance mechanisms in the case of pyrethroid-PBO nets). These should be articulated in the registry a priori and in the report.

6. Future considerations

It should be noted that the data from experimental hut studies used to assess the non-inferiority of ITNs provide limited insight into the bioefficacy of the treatment over time in the field or whether personal protective efficacy is maintained when nets are subjected to normal wear and tear. The testing of washed (20 times) ITNs uses the wash durability of the ITN as a surrogate for natural ageing, but the relationship between wash resistance and field durability remains unclear *(6)*. This is a clear limitation of the approach used to determine the non-inferiority of candidate ITNs, which may require further consideration when data on the relationship between 20 times washed and operationally aged ITNs become available.

Techniques other than experimental huts may be equally suitable for non-inferiority assessments and may offer certain advantages. Experimental hut trials with free-flying mosquitoes are currently the preferred method for evaluating vector control products designed to be used indoors. However, infrastructure requirements mean that these tests can presently only be carried out in a small number of sites, mainly in Africa. Consequently, products can only be evaluated against a limited number of mosquito vector populations. Furthermore, experimental hut trials' reliance on having a sufficient number of local free-flying mosquitoes means that study duration is affected by the season and level of routine local mosquito control. Resources permitting, other potential alternative testing methods, e.g. the I-ACT ambient chamber test or the tunnel test, may be used for non-inferiority studies alongside experimental huts in order to investigate whether these other methods present suitable alternatives for generating non-inferiority data.

References

1. Norms, standards and processes underpinning development of WHO recommendations on vector control. Geneva: World Health Organization; 2020 (https://iris.who.int/handle/10665/338030, accessed 10 October 2023).
2. WHO handbook for guideline development, second edition. Geneva: World Health Organization; 2014 (https://iris.who.int/handle/10665/145714, accessed 10 October 2023).

3. Guidelines Review Committee [website]. In: Groups. Geneva: World Health Organization (https://www.who.int/groups/guidelines-review-committee, accessed 10 October 2023).
4. WHO malaria policy advisory committee meeting: meeting report, October 2017. Geneva: World Health Organization; 2017 (https://iris.who.int/handle/10665/259396, accessed 10 October 2023).
5. Notice of intent to introduce non-inferiority studies as a component of the vector control evaluation process. Geneva: World Health Organization; 2018.
6. Data requirements and protocol for determining non-inferiority of insecticide-treated net and indoor residual spraying products within an established WHO policy class. Geneva: World Health Organization; 2019 (https://iris.who.int/handle/10665/276039, accessed 10 October 2023).
7. Technical consultation on determining non-inferiority of vector control products within an established class: report of a virtual meeting, 31 August–2 September 2021. Geneva: World Health Organization; 2021 (https://iris.who.int/handle/10665/349446, accessed 10 October 2023).
8. Indoor residual surface treatments for malaria transmission control in areas with insecticide-resistant mosquito populations: preferred product characteristics. Geneva: World Health Organization; 2022 (https://iris.who.int/handle/10665/356901, accessed 10 October 2023).
9. WHO malaria policy advisory group (MPAG) meeting: meeting report, October 2021. Geneva: World Health Organization; 2021 (https://iris.who.int/handle/10665/348220, accessed 10 October 2023).
10. WHO malaria policy advisory group (MPAG) meeting: meeting report, October 2022. Geneva: World Health Organization; 2022 (https://iris.who.int/handle/10665/364715, accessed 10 October 2023).
11. WHO malaria policy advisory group (MPAG) meeting report, 18–20 April 2023. Geneva: World Health Organization; 2023 (https://iris.who.int/handle/10665/368391, accessed 10 October 2023).
12. Technical consultation to assess comparative efficacy of vector control products: meeting report, 5 and 9 June 2023. Geneva: World Health Organization; 2023 (https://iris.who.int/handle/10665/372841, accessed 10 October 2023).
13. WHO guidelines for malaria, 14 March 2023. Geneva: World Health Organization; 2023 (https://iris.who.int/handle/10665/366432, accessed 10 October 2023).
14. Sherrard-Smith E, Ngufor C, Sanou A, Guelbeogo MW, N'Guessan R, Elobolobo E, et al. Inferring the epidemiological benefit of indoor vector control interventions against malaria from mosquito data. Nat Commun. 2022;13(1):3862. doi:10.1038/s41467-022-30700-1.
15. World Health Organization, WHO Pesticide Evaluation Scheme. Guidelines for laboratory and field testing of long-lasting insecticidal nets. Geneva: World Health Organization; 2013 (WHO/HTM/NTD/WHOPES/2013.1; https://iris.who.int/handle/10665/80270, accessed 10 October 2023).
16. Guidelines for testing mosquito adulticides for indoor residual spraying and treatment of mosquito nets. Geneva: World Health Organization; 2006 (WHO/CDS/NTD/WHOPES/GCDPP/2006.3; https://iris.who.int/handle/10665/69296, accessed 10 October 2023).
17. Everson-Stewart S, Emerson SS. Bio-creep in non-inferiority clinical trials. Stat Med. 2010;29:2769–80. doi:10.1002/sim.4053.
18. Challenger JD, Nash RK, Ngufor C, Sanou A, Toé KH, Moore S, et al. Assessing the variability in experimental hut trials evaluating insecticide-treated nets against malaria vectors. Curr Res Parasitol Vector Borne Dis. 2023;3:100115. doi:10.1016/j.crpvbd.2023.100115.

19. Manual for monitoring insecticide resistance in mosquito vectors and selecting appropriate interventions. Geneva: World Health Organization; 2022 (https://iris.who.int/handle/10665/356964, accessed 10 October 2023).
20. Gimnig JE, Walker ED, Otieno P, Kosgei J, Olang G, Ombok M, et al. Incidence of malaria among mosquito collectors conducting human landing catches in western Kenya. Am J Trop Med Hyg. 2013;88(2):301–8. doi:10.4269/ajtmh.2012.12-0209.
21. Report of the twentieth WHOPES working group meeting, WHO/HQ, Geneva, 20–24 March 2017. Geneva: World Health Organization; 2017 (WHO/HTM/NTD/WHOPES/2017.04; https://iris.who.int/handle/10665/258921, accessed 10 October 2023).
22. Indoor residual spraying: an operational manual for indoor residual spraying (IRS) for malaria transmission control and elimination, second edition. Geneva: World Health Organization; 2015 (https://iris.who.int/handle/10665/177242, accessed 10 October 2023).
23. Lissenden N, Armistead JS, Gleave K, Irish SR, Martin JL, Messenger LA, et al. Developing consensus standard operating procedures (SOPs) to evaluate new types of insecticide-treated nets. Insects. 2021;13(1):7. doi:10.3390/insects13010007.
24. Ngufor C, Critchley J, Fagbohoun J, N'Guessan R, Tdijnou D, Rowland M. Chlorfenapyr (a pyrrole insecticide) applied alone or as a mixture with alpha-cypermethrin for indoor residual spraying against pyrethroid resistant *Anopheles gambiae sl*: an experimental hut study in Cove, Benin. PLoS One. 2016;11(9):e0162210. doi:10.1371/journal.pone.0162210.
25. CIPAC handbooks [website]. Milton Keynes: Collaborative International Pesticides Analytical Council (https://www.cipac.org/index.php/m-p/handbooks, accessed 10 October 2023).

Annex 1. List of active comparators (October 2023)

Chemical class	Product type*	Standard of care	1. FIC product	2. SIC product with evidence of public health benefit available from RCTs	3. SIC product that has shown superiority to the FIC product on the primary end-point in entomological trials	4. Best performing product among the SIC products
Sodium channel modulators (pyrethroids)	ITN	N/A	Not applicable to ITNs or IRS, may become applicable to new vector control products			
Sodium channel modulators (pyrethroids) with synergist PBO	ITN	Currently pyrethroid-only ITN	Olyset™ Plus	PermaNet 3.0		TSARA Boost Veeralin® Yorkool G3® DuraNet Plus©
Uncouplers of oxidative phosphorylation via disruption of the proton gradient (pyrroles)	ITN	Currently pyrethroid-only ITN	Interceptor® G2			PermaNet® Dual
Juvenile hormone mimics (pyriproxyfen)	ITN	Currently pyrethroid-only ITN	Royal Guard®			
Acetylcholinesterase (AChE) inhibitors (carbamates, organophosphates)	IRS	Negative control / IRS currently used in the area	Actellic® Ficam®			
Nicotinic acetylcholine receptor competitive modulators (neonicotinoids)	IRS	Negative control / IRS currently used in the area	SumiShield®			
GABA-gated chloride channel allosteric modulators (meta-diamides and isoxazolines)	IRS	Negative control / IRS currently used in the area	Vectron™ T500			

*Covers products currently with a WHO recommendation – new products will be added as they are recommended.

Annex 2. Data template

See Excel file in Dropbox folder.

Annex 3. Data analysis code

This tutorial is designed to provide guidance on carrying out non-inferiority assessments for insecticide-treated products in experimental hut trials. The methodology used here matches that used by the WHO Global Malaria Programme for assessing the comparative efficacy of vector control products *(1).* In these trials, volunteers sleep inside the huts with the vector control products, and mosquitoes are collected early the next morning to gather information on mosquito mortality and mosquito blood feeding. Both the products and the volunteers are rotated through the huts to avoid potential biases.

Once the data have been collected, they are analysed using logistic regression models, with separate models fitted for mosquito mortality and mosquito blood feeding. Fixed effects should be included for the treatment arm, hut, sleeper and day of the trial. For ITNs, it is common to include unwashed and washed nets of the same brand in the same trial. Here, washing the nets is designed to reproduce the effects of ageing on the nets. In this way, one can assess the longevity of the product. For ITNs, therefore, one can conduct separate non-inferiority assessments for unwashed and washed nets, as well as a combined assessment in which unwashed and washed nets of the same brand are assessed together. The fixed effects included in the logistic regression are brand of net, hut, sleeper, day of the trial and whether the net was washed or not.

The data set used in this tutorial to illustrate the methodology is a synthetic dataset (that is, one generated by computer simulation), rather than one from a real-world experimental hut trial. The treatment arms are different types of bed nets, but the same methodology can be used for IRS. The treatment arms in the data set are named for their role in the non-inferiority assessment. The treatment arms are: an untreated control net, a standard comparator (unwashed and washed), an active comparator (unwashed and washed), and a candidate net (unwashed and washed). Therefore, there are seven treatment arms in total. The candidate net should be of the same chemistry as the active comparator. For example, in the context of ITNs, this could be a dual-AI net, such as a pyrethroid-PBO net. The active comparator (sometimes referred to as the FIC product) should be a prequalified product (e.g. one that has shown significant efficacy in a RCT). In these trials, the aim is to assess the non-inferiority of the candidate net compared to the active comparator. As this assessment does not provide information on product efficacy in absolute terms, the evaluation will also look at whether the candidate net is superior to the standard comparator. In the context of ITNs, the standard comparator is usually a pyrethroid-only net.

Finally, a note on the choice of the non-inferiority margin used here: The efficacies of the candidate net and the active comparator (for mosquito mortality or blood-feeding inhibition) are compared by constructing an OR. This OR and its 95% CI should then be compared to the non-inferiority margin. For mosquito mortality, the entire 95% CI must lie **above** the non-inferiority margin for the candidate product to be non-inferior to the active comparator. If this is not the case, the candidate is said to be "not non-inferior" to the active comparator. By contrast, for blood feeding, the entire 95% CI must lie **below** the non-inferiority margin for non-inferiority to be achieved.

In this work, the non-inferiority margin is set so that the candidate net efficacy is no more than 7% lower than that of the active comparator. Therefore, when assessing mosquito mortality, the non-inferiority margin is chosen so that the mosquito mortality measured for the candidate net is no more than 7% lower than that measured for the active comparator in order for non-inferiority to be achieved. This means that the OR for the non-inferiority margin will vary from trial to trial and must be calculated for each assessment.

Click here for the tutorial on non-inferiority assessments for experimental hut trials.

Reference

1. Technical consultation on determining non-inferiority of vector control products within an established class: report of a virtual meeting, 31 August–2 September 2021. Geneva: World Health Organization; 2021 (https://iris.who.int/handle/10665/349446, accessed 10 October 2023).

7. Annex 4. Data checklist

Section / topic	Item number	Checklist item
Title and abstract		
	1a	Structured summary of trial design, methods, results and conclusions
	1b	Test items (test products), test system (mosquito species), study site, resistance profile of test system
Introduction		
Background and objectives	2a	Specific objectives or hypotheses
	2b	Outcomes measured (primary and secondary)
Methods		
Trial design	3a	Description of trial design, blinding and randomization, number of huts / replicates
	3b	Important changes to methods after trial commencement (such as deviations or changes to sample size), with reasons
Test system	4a	Study arms including test item preparation (detailed washing, application or spraying procedure)
	4b	Settings and locations where the data were collected
	4c	Characteristics of local vector population (species, phenotypic resistance, i.e. resistance ratio or WHO tube test results ideally conducted during the study, resistance mechanisms)
	4d	Characteristics of mosquitoes used in quality assurance testing. Report mortality at WHO discriminating doses (conducted during study), resistance ratio and resistance phenotype, resistance mechanisms
Interventions	5	The interventions for each group with sufficient detail to allow for study replication, including storage conditions, transfer and procedures, i.e. washing, holing, hanging, storage for ITNs, and chemical analysis of AIs (including synergists) before and after test item preparation and after completion of the trial, spraying procedures and insecticide quantification for IRS
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
	6b	Proportional mosquito mortality Proportional mosquito fertility (vector control products with fertility effects)
	6c	Proportional mosquito blood feeding Proportional mosquito mortality (vector control products with fertility effects)
	6d	Duration of efficacy: 0 and 20 washes for ITNs Months of efficacy (> 80% mortality) for IRS

	6e	Environmental conditions during testing
	6f	Any changes to trial outcomes after the trial commenced, with reasons
Sample size	7	How sample size was determined, including mosquito densities, variability for hut, volunteer and day used for simulations, mortality of active comparator used and margin of non-inferiority
Randomization		
Sequence generation	8a	Method used to generate the random allocation sequence
	8b	Type of randomization
Allocation concealment mechanism	9a	Mechanism used to implement the random allocation sequence (such as lottery / opaque envelopes, number generator), describing any steps taken to conceal the sequence until interventions were assigned
	9b	Include the rotation scheme in an appendix
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions
Blinding	11a	If done, who was blinded after assignment to interventions (e.g. participants, those conducting the trial, those analysing the data) and how
	11b	If relevant, description of blinding of interventions (e.g. size, shape and numbering of nets or sachets for IRS, codes used to conceal allocation)
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes using non-inferiority, accounting for sources of variability and bias including huts, sleepers and day
	12b	Methods for secondary analyses, including superiority analyses for active comparator and candidate relative to control or standard of care
Results		
Primary outcome(s)	13a	For ITNs, non-inferiority of candidate (test product) compared to active comparator using proportional mosquito mortality (24-hour mortality unless there is justification for longer holding period up to 168 hours) and in the case of ITNs with juvenile hormone analogue reproductive inhibition measured by OR, or eggs laid measured by RR For IRS, non-inferiority of candidate (test product) compared to active comparator using proportional mosquito mortality measured by OR (24-hour mortality, unless there is justification for a longer holding period up to 168 hours)
Secondary outcomes	13b	Duration of mortality for IRS or efficacy of washed versus unwashed nets (added as a fixed effect in secondary analysis)
	13c	For ITNs, non-inferiority of candidate (test product) compared to active comparator using proportional mosquito blood feeding In the case of ITNs with juvenile hormone analogue reproductive inhibition, mosquito mortality at 72 hours measured by OR
	13d	A table showing primary and secondary outcomes using an appropriate measure of centrality and dispersion: i) total mosquitoes; ii) mosquito number per night (e.g. arithmetic mean, geometric mean, Williams mean with 95% CI or median with range); iii) total

		mosquitoes dead; iv) arithmetic mean % (95% CI) mortality at each holding time; v) total mosquitoes fed; vi) arithmetic mean % (95% CI) blood-feeding rate All outcomes should be summarized for the study as a whole and reported per arm; blood-feeding inhibition may also be reported
Numbers analysed	13e	For each group, number of mosquitoes (denominator) included in each analysis
	13f	Number of replicates (huts, sleepers, interventions, nights of collection, duration of study)
Outcomes and estimation	13g	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% CI)
	13h	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	14	Results of any other data collected as part of trial quality assurance, including laboratory tests such as cone bioassay and tunnel tests, as well as any chemical verification of AI. All tests must report the absolute number of mosquitoes tested (e.g. the discriminating dose bioassay must report the number of mosquitoes tested and the number that died for each replicate).
Harms	15	All important harms or unintended effects in each group (adverse events)
Discussion		
Limitations	16	Trial limitations, addressing sources of potential bias, imprecision and, if relevant, multiplicity of analyses
Generalizability	17	Generalizability of the trial findings
Interpretation	18	Interpretation consistent with results, considering other relevant evidence, e.g. cone bioassays, resistance ratio of wild mosquito populations
Other information		
Registration	19	Registration number of study
Protocol	20	Where the full trial protocol can be accessed
Funding	21	Sources of funding and other support (such as supply of active comparators), role of funders
Appendices		
All raw data in standard Excel spreadsheet (using template in Annex 2)		

A tutorial on non-inferiority assessments for experimental hut trials

Joseph D. Challenger & The Global Malaria Programme, World Health Organisation

August 2023

Contents

Chapter 1

Introduction

This tutorial is designed to provide guidance on carrying out non-inferiority assessments for insecticide-treated products in experimental huts trials (EHTs). The methodology used here matches that used by the Global Malaria Programme for assessing the comparative effectiveness of vector control products[1]. In these trials, volunteers sleep inside the huts with the vector control products, and mosquitoes are collected early the next morning, to gather information on mosquito mortality and mosquito blood feeding. Both the products and the volunteers are rotated around the huts, to avoid potential biases.

Once the data have been collected, they are analysed using logistic regression models, with separate models fitted for mosquito mortality and mosquito blood feeding. Fixed effects should be included for the treatment arm, hut, sleeper, and the day of the trial. For insecticide-treated nets (ITNs), it is common to include unwashed and washed nets of the same brand in the same trial. Here, washing the nets is designed to reproduced the effects of ageing. In this way one can assess the longevity of the product. For ITNs, therefore, one can make separate non-inferiority assessments for unwashed and washed nets, as well as a combined assessment, where unwashed and washed nets of the same brand are assessed together. For this analysis, the fixed effects included in the logistic regression are: brand of net, hut, sleeper, the day of the trial, and whether the net has been washed or not.

The dataset used in this tutorial to illustrate the methodology, is a synthetic dataset (that is, one generated from computer simulation), rather than one from a real-world EHT. The treatments arms are different types of bed net, but the same methodology can be used for *e.g.* indoor residual spraying (IRS). The treatment arms in the dataset are named for their role in the non-inferiority trial. The treatment arms are: an untreated control net, a standard comparator (unwashed and washed), an active comparator (unwashed and washed), and a candidate net (unwashed and washed). Hence, in total there are seven treatment arms. The candidate net should be of the same chemistry as the active comparator. For example, in the context of ITNs, this could be a dual-chemistry, such as pyrethroid and piperonyl-butoxide (PBO). The active comparator (sometimes referred to as the 'first in class product') should be a pre-approved product (*e.g.* one that has shown significant efficacy in a cluster randomised trial). In these trials, we wish to assess the non-inferiority of the candidate net compared to the active comparator. As this assessment does not provide information on the products efficacy in absolute terms, we will also test whether the candidate net is superior to the standard comparator. In the context of ITNs, the standard comparator is usually a pyrethroid-only net.

[1]See *Technical consultation on determining non-inferiority of vector control products within an established class* Report of a virtual meeting 31 August–2 September 2021, World Health Organisation

Finally, we should discuss the choice of the non-inferiority margin (NIM) used here. The efficacy of the candidate net and the active comparator (be it for mosquito mortality or blood feeding inhibition) are compared by constructing an odds ratio (OR). This OR and its 95% confidence interval should then be compared to the NIM. For mosquito mortality, the entire 95% confidence interval must lie *above* the NIM for the candidate product to be non-inferior to the active comparator. If this is not the case, the candidate is said to be 'not non-inferior' to the active comparator. In contrast, for blood feeding the entire 95% confidence interval must lie *below* the NIM for non-inferiority to be achieved.

In this work, the NIM is set so that the candidate net efficacy should be no more than 7% lower than that of the active comparator. So, when assessing mosquito mortality, the NIM is chosen so that mosquito mortality measured for the candidate net must be no more than 7% lower than that measured for the active comparator, in order for non-inferiority to be achieved. This means that the OR for the NIM will vary from trial to trial, and must be calculated for each assessment. We will show examples of this in the following chapters. Chapter 2 will outline the procedure in R; Chapter 3 will follow the same procedure using STATA.

Chapter 2

Analysis in R

A brief introduction to R

The code included in this tutorial is written in R. We do not include a comprehensive introduction to R, as many others are available elsewhere. However, we will include a few brief comments on the syntax of R. Variables (be they numbers, or character strings) can be stored in the internal memory using either `=` or `<-`. For example, writing `x<-5` assigns a value of 5 to `x`. The symbol `#` is used to indicate a comment. That is, anything that follows a `#` will not be read as R code.

R contains a number of core functions, which carry out commonly used operations. Additional functions can be found within *packages*. A package can be loaded using the `library()` function. For example, we can load the ggplot2 package, which is a versatile library for making graphs, by running the following command:

```
> library(ggplot2)
```

If you have not used this package before, you may need to download it. You can do this by running the command `install.packages('ggplot2')`. Alternatively, if you're using Rstudio, you can click on the 'Tools' menu, then click on 'Install packages...', and search for the desired package.

2.1 Loading & summarising a dataset

To demonstrate the statistical analyses to be carried out, we will use a simulated datasetThis dataset has been uploaded with these materials, along with an R script containing the work outlined in this tutorial. To run the R script you should download R & RStudio. If you wish to run the script, you should download the ZIP file, and extract the folder. Then, double-click on the project file `R_tutorial.Rproj` to open it in RStudio. We recommend doing this, as this should mean that the R session will open with the current working directory being the same as the location of our files. This will make it easier for R to find all our code.

Once the R project file is open, you can then open the R script `R_tutorial.R`. First we will load the packages that we will use:

```
library(ggplot2)
library(lme4)
library(cowplot)
```

Running the command `source('useful_functions_tutorial.R')` will load some user-defined functions that are stored in another file within this project. If you wish, you can open this file to inspect the functions.

The dataset has been stored as an `.csv` file, and can be loaded using the following command:

```
> df <- read.csv('example_dataset.csv')
```

This stores the data in the variable `df`. Now we can see the contents of this dataset, using the `str()` function.

```
> str(df)
'data.frame': 343 obs. of  14 variables:
 $ day      : int  1 1 1 1 1 1 1 2 2 2 ...
 $ hut      : int  1 2 3 4 5 6 7 1 2 3 ...
 $ sleeper  : int  2 3 4 5 6 7 1 3 4 5 ...
 $ treatment: chr  "Control" "Standard_comparator_unwashed" ...
 $ ITN      : chr  "Control" "Standard_comparator" ...
 $ wash     : int  0 0 1 0 1 0 1 0 0 1 ...
 $ unf_live : int  7 6 1 13 12 3 9 4 11 7 ...
 $ unf_dead : int  0 1 1 8 3 1 2 0 0 0 ...
 $ bf_live  : int  6 1 1 7 3 3 4 1 1 2 ...
 $ bf_dead  : int  0 1 0 2 2 1 1 0 0 0 ...
 $ tot_dead : int  0 2 1 10 5 2 3 0 0 0 ...
 $ tot_bf   : int  6 2 1 9 5 4 5 1 1 2 ...
 $ total    : int  13 9 3 30 20 8 16 5 12 9 ...
```

Here we see that the dataset contains 343 data points. This trial has 7 arms and runs over 7 weeks (49 days). Let's look at this in more detail. The variable `treatment` defines which trial arm each data point relates to. We can summarise the trial arms like this:

```
> table(df$treatment)

  Active_comparator_unwashed     Active_comparator_washed
                          49                           49
         Candidate_unwashed             Candidate_washed
                          49                           49
                     Control Standard_comparator_unwashed
                          49                           49
  Standard_comparator_washed
                          49
```

Hence we can see that the trial contains 3 ITNs: a standard comparator, an active comparator, and the candidate net. For each ITN, there are 2 trial arms, containing unwashed and washed nets. Additionally, the trial contains an arm with an untreated control net. The dataset also contains an alternative way to describe the trial arms, using the variables `ITN` and `wash`. The latter variable takes a value 0 (unwashed) or 1 (washed). You can look at a summary of these variables running `df$ITN` and `df$wash`.

The variable `total` records the total number of mosquitoes collected in a given hut on a given night. These mosquito counts are broken down to indicate whether or not each mosquito has died

or blood fed: `unf_live` = unfed & alive; `unf_dead` = unfed & dead; `bf_live` = blood fed & live; `bf_dead` = blood fed & dead. The dataset also summarises the total number of dead mosquitoes (`tot_dead`), and the total number of blood-fed mosquitoes (`tot_bf`).

We can also see that the study contains 7 huts and 7 sleepers:

```
> table(df$hut)
 1  2  3  4  5  6  7
49 49 49 49 49 49 49

> table(df$sleeper)
 1  2  3  4  5  6  7
49 49 49 49 49 49 49
```

It will be useful to change the data types of some of the variables. For the terms that we will include as fixed effects in the model, we will make these *factor variables* in R:

```
df$hut <- as.factor(df$hut)
df$sleeper <- as.factor(df$sleeper)
df$day <- as.factor(df$day)
df$treatment <- as.factor(df$treatment)
df$ITN <- as.factor(df$ITN)
```

This is particularly important for variables like `hut` and `sleeper`, where the numbering is purely to label the huts. We want our regression model to recognise that each number indicates a distinct category: it does not represent a quantitative measurement. Finally for this section, we will look at a summary of the mosquito mortality and blood-feeding across the 7 trial arms. These summaries are taken directly from the dataset (*i.e.* they are not adjusted estimates derived from a fitted regression model), using the user-defined function `summm()`. See Appendix A for a full definition of this function.

```
> tab_mortality <- summm(df, vec = df$treatment, td = 'tot_dead',
                           tot = 'total', table = 1)
> tab_mortality
                          Arm Percentage
1                     Control       5.99
2 Standard_comparator_unwashed      11.61
3   Standard_comparator_washed       8.75
4   Active_comparator_unwashed      34.09
5     Active_comparator_washed      21.45
6           Candidate_unwashed      28.99
7             Candidate_washed      22.54
> tab_bf <- summm(df, vec = df$treatment, td = 'tot_bf', tot = 'total', table = 1)
> tab_bf
                          Arm Percentage
1                     Control      33.81
2 Standard_comparator_unwashed      17.41
3   Standard_comparator_washed      23.96
4   Active_comparator_unwashed      17.05
5     Active_comparator_washed      25.81
6           Candidate_unwashed      16.77
7             Candidate_washed      26.16
```

2.2 Making the non-inferiority assessment

For each hut trial of ITNs, we will make 6 separate non-inferiority assessments. Note that, typically, the blood-feeding assessments are not carried out for trials of IRS. For ITNs, the 6 assessments are as follows:

1. Mosquito mortality: comparing the unwashed candidate to the unwashed active comparator
2. Mosquito mortality: comparing the washed candidate to the washed active comparator
3. Mosquito mortality: comparing the candidate to the active comparator (unwashed and washed combined)
4. Blood feeding: comparing the unwashed candidate to the unwashed active comparator
5. Blood feeding: comparing the washed candidate to the washed active comparator
6. Blood feeding: comparing the candidate to the active comparator (unwashed and washed combined)

For each individual assessment, we can break it down into the following steps:

- Calculate the unadjusted mosquito mortality (or blood feeding) directly from the data (*i.e.* without fitting a regression model)
- Choose which trial arm should be used as the baseline category in the regression model (here it should be the active comparator). Then fit the regression model, adjusting for hut, sleeper and day.
- Calculate the odds ratio for mosquito mortality in the standard comparator, compared to that observed in the active comparator arm.
- Calculate the non-inferiority margin to use, based on the unadjusted mosquito mortality observed in the active comparator arm. Then make the non-inferiority assessment.
- Now check that the candidate net is superior to the standard comparator. First we change the baseline category of the regression model, then we fit the same regression model as before to the data. The p-value for the fixed effect will be used to test for superiority.

In the R code, we also provide code to make visualisations of the non-inferiority assessment, which we shall also outline here. We shall now make a brief comment about the type of regression models we will use here– the logistic regression model. This type of model is used to model data that can be described as proportions (*e.g.* proportion of mosquitoes killed, or proportion of mosquitoes blood fed). To fit this model, the proportion in question, often denoted as p, is transformed onto the log-odds scale. If we define the transformed value as $X(p)$, we can write:

$$X(p) = \log\left(\frac{p}{1-p}\right).$$

Note that, on the log-odds scale, values can be positive or negative. A value of 0 corresponds to $p = 0.5$. We can write the inverse function like this:

$$p = \text{InvLogit}(X) = \frac{\exp(X)}{\exp(X)+1}.$$

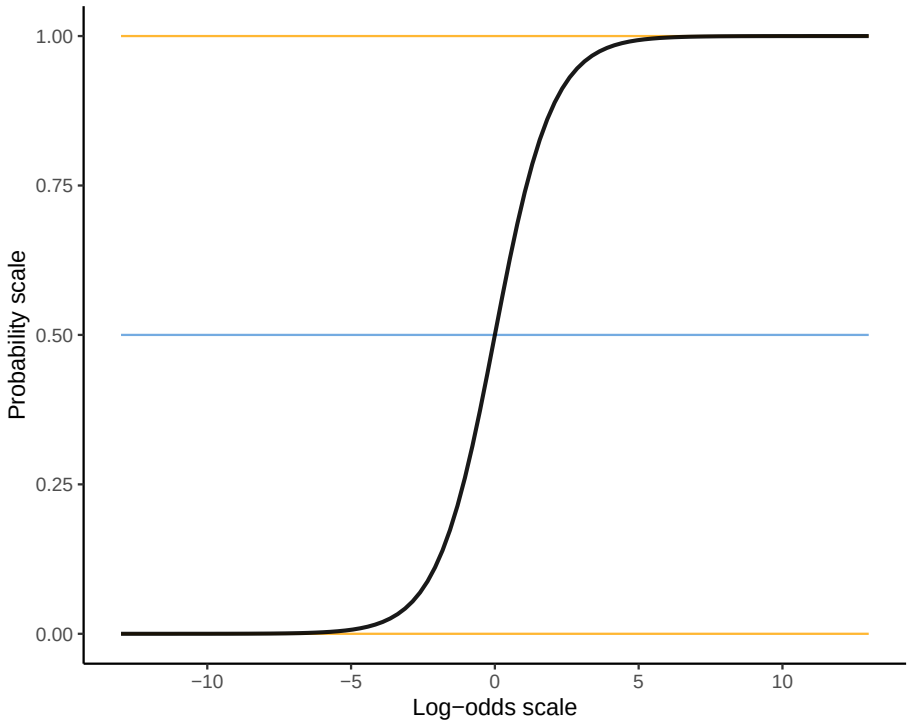


Figure 2.1: Relationship between the log-odds and probability scales. The function is symmetric about 0, which corresponds to a probability of 0.5. The probability approaches 1 as the log-odds value becomes very large (technically, we say 'as it tends to infinity'). Similarly, the probability goes to 0 as the log-odds value tends to minus infinity. The blue horizontal line locates $p = 0.5$, which corresponds to a value of 0 on the log-odds scale.

Whilst a proportion is restricted to the interval [0,1], $X(p)$, can take any value, positive or negative. This facilitates the fitting of the regression model: the model is not hampered by a lower limit as p approaches 0, or an upper limit as p approaches 1. Figure 2.1 shows this transformation.

2.2.1 Mosquito mortality

Let's start with the first non-inferiority assessment listed above (mosquito mortality, unwashed nets). We will fit a regression model with the following form:

```
fit1 <-
  glm(
    cbind(tot_dead, total - tot_dead) ~
      treatment + hut + sleeper + day,
    family = binomial, data = df)
```

Let's define the terms mentioned here. The function `glm()`, which comes from the `lme4` package will fit a generalised linear regression model to the data (we use the argument `data` to tell the function which data to use). Setting `family = binomial` indicates that we wish to fit a logistic regression model. We have asked that the model output be stored in the container `fit1`. Note that the numbering of the fitted models in this tutorial (`fit1`, `fit2`, *etc.*) matches with the 6 non-inferiority assessments listed at the beginning of this section.

The first argument included inside the function `glm()`, `cbind(tot_dead, total - tot_dead)`, indicates that the mosquito counts can be split into two categories: dead (`tot_dead`), or alive (the number of alive mosquitoes can be written as the total number of mosquitoes minus the number of dead mosquitoes). All terms after the tilde symbol (`~`) are terms we wish to include in the model. Here we include fixed effects for the treatment arm, hut, sleeper and day of the trial. Before proceeding further, we should think about how the model will interpret these fixed effects: we will use `treatment` as an example, but the same principles hold to the other variables too. As we saw earlier, `treatment` can take 7 different values- one for each arm in the trial. When fitting the model, the function `glm()` will choose one of these 7 values as a 'baseline' category (by default, the first category in alphabetical or numerical value will be chosen). Then the value of 6 parameters will be estimated; these can be thought of as *offsets*, differences between a given trial arm and the baseline trial arm. The choice of which trial arm is an arbitrary one, and does not affect the non-inferiority assessment. However, the assessment is made simpler if the active comparator is used as the baseline category. Then, one can simply read off the parameter value associated with the candidate net: this tells us the difference in performance of the candidate net, compared to the active comparator. If the treatment variable is a factor variable in R, `glm()` will set the lowest factor to be the baseline category. We can check the ordering of factors using the `levels()` function:

```
> levels(df$treatment)
[1] "Active_comparator_unwashed"   "Active_comparator_washed"
[3] "Candidate_unwashed"           "Candidate_washed"
[5] "Control"                      "Standard_comparator_unwashed"
[7] "Standard_comparator_washed"
```

Here we can see that `"Active_comparator_unwashed"` is the lowest level in the factor (this is because the levels are ordered alphabetically by default). So, this means we don't have to reorder the levels, but later we'll meet an example where we do need to do this. We can now fit the model defined above (`fit1`). When we have done this, we can use the `summary()` function to view the fitted model

```
> summary(fit1)

Coefficients:
                                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                         -0.0690133  0.3184980  -0.217 0.828455
treatmentActive_comparator_washed   -0.5350687  0.1545003  -3.463 0.000534 ***
treatmentCandidate_unwashed         -0.0286659  0.1567682  -0.183 0.854911
treatmentCandidate_washed           -0.5543215  0.1655243  -3.349 0.000811 ***
treatmentControl                    -2.0162343  0.2021793  -9.973  < 2e-16 ***
treatmentStandard_comparator_unwashed -1.2311743  0.1855108  -6.637 3.21e-11 ***
treatmentStandard_comparator_washed -1.6727303  0.2036198  -8.215  < 2e-16 ***
hut2                                -0.3197885  0.1578667  -2.026 0.042797 *
hut3                                -0.4449210  0.1603844  -2.774 0.005536 **
hut4                                -1.0380668  0.1867768  -5.558 2.73e-08 ***
hut5                                -0.9908420  0.1783093  -5.557 2.75e-08 ***
hut6                                -1.1492457  0.1921662  -5.980 2.22e-09 ***
hut7                                -1.4119923  0.1844370  -7.656 1.92e-14 ***
sleeper2                            -0.1441138  0.1894002  -0.761 0.446719
```

```
sleeper3                          0.1593957  0.1906211   0.836 0.403047
sleeper4                          0.0880827  0.1966041   0.448 0.654138
sleeper5                          0.5207049  0.1795870   2.899 0.003738 **
sleeper6                          0.3354179  0.1861006   1.802 0.071491 .
sleeper7                          0.5837013  0.1763917   3.309 0.000936 ***
day2                             -0.7682702  0.5383765  -1.427 0.153576
day3                             -0.2172527  0.4005852  -0.542 0.587585
day4                              0.3238629  0.4541270   0.713 0.475750
day5                             -0.3809665  0.4612852  -0.826 0.408872
day6                              0.3210389  0.4113972   0.780 0.435178
...
day47                            -0.4420176  0.4107280  -1.076 0.281846
day48                             0.3312319  0.3657527   0.906 0.365139
day49                            -0.5292627  0.4886953  -1.083 0.278803
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Here, the output has been truncated: in particular, we've dropped the fixed-effect parameters for days 4-46 (inclusive), for conciseness. The intercept category gives the log-odds-transformed proportion of mosquitoes killed in the trial arm with unwashed active comparator ITNs, used in hut 1, with sleeper 1, on day 1. This value, -0.069, corresponds to a proportion of ≈ 0.485 (*i.e.* nearly half of the mosquitoes killed). Let's now calculate the odds-ratio between the unwashed candidate and the unwashed active comparator nets. The OR is simply the exponent of the offset parameter for the unwashed candidate net (the estimate of this parameter is given above as -0.0286...). To calculate this, we need to extract the desired parameter from the fitted model (`fit1`). This can be done using the row & column numbers of the table of coefficents. Here, we shall chose the slightly more verbose option of identifying the parameters by name, like this:

```
>coef(summary(fit1))[‘treatmentCandidate_unwashed’,"Estimate"]
-0.02866594
```

Check that you can see where this value has come from in the model output. To calculate the OR, we simply take the exponent of this value. We'll store this as `OR1`, as it is associated with model 1.

```
>OR1 <- exp(coef(summary(fit1))[‘treatmentCandidate_unwashed’,"Estimate"])
```

In this example, `OR1=0.9717...`. We can use the standard error of this parameter estimate to form the 95% confidence interval like this:

```
>OR1_lower <- exp(coef(summary(fit1))[‘treatmentCandidate_unwashed’,‘Estimate’] -
              1.96*coef(summary(fit1))[‘treatmentCandidate_unwashed’,‘Std. Error’])
>OR1_upper <- exp(coef(summary(fit1))[‘treatmentCandidate_unwashed’,‘Estimate’] +
              1.96*coef(summary(fit1))[‘treatmentCandidate_unwashed’,‘Std. Error’])
```

In order to carry out the non-inferiority assessment, we need to set the non-inferiority margin (NIM). In these guidelines, we use a variable NIM, which depends on the level of mosquito mortality (or blood-feeding) observed in the first-in-class product (the active comparator). This is because using a fixed NIM can make it difficult to show non-inferiority when the first-in-class product is highly efficacious (*e.g.* if the mosquito mortality is >80%). Here the NIM is defined so that the proportion of mosquitoes killed by the candidate net should be no more than 7% less than the proportion killed by the active comparator. To determine the NIM, we will use the unadjusted estimates of mosquito mortality, that we calculated in Section 1:

```
FIC_mortality1 <- tab_mortality[tab_mortality$Arm==‘Active_comparator_unwashed’,
    ]$Percentage / 100
# Calculate the odds ratio (OR) for a mortality 7% lower than this one:
non_inf_margin1 <- ((FIC_mortality1 - 0.07) / (1- (FIC_mortality1 - 0.07))) /
    (FIC_mortality1 / (1- FIC_mortality1))
```

This gives a NIM of 0.7183... . We have developed a user-defined function to make the non-inferiority assessment and produce a visualisation of it. The function is called `plot_NI_OR()`, and must be provided with the OR and the NIM (see Appendix A for full function definition):

```
plot_NI_OR(OR = OR1, ORl = OR1_lower, ORu = OR1_upper, mortality = 1,
        NIM = non_inf_margin1, precision = 3)
```

Calling this function, will also lead to the outcome of the non-inferiority assessment being printed to the console. In this case, it produces this:

```
[1] "OR=0.972 [0.715, 1.321]"
[1] "NOT non-inferior"
```

In this case, the lower confidence interval (0.714...) is slightly less than the NIM (0.7183). So we cannot say that the unwashed candidate net is non-inferior to the unwashed active comparator. This assessment is summarised in Figure 2.2B: however, as the 95% CI only just passes below the NIM, the outcome is not so clear here. Therefore, we must use the (unrounded) numerical values for the 95% CI of the OR and the NIM to make the assessment.

In addition to the non-inferiority assessment, we must also check that the unwashed candidate net is superior to the unwashed standard comparator. We can do this using the same regression model used above, but it will be more straightforward if we change the baseline category first, using the `relevel()` function:

```
df$treatment <- relevel(df$treatment, ‘Standard_comparator_unwashed’)
> levels(df$treatment) # check the levels of the factor
[1] "Standard_comparator_unwashed" "Active_comparator_unwashed"
[3] "Active_comparator_washed"     "Candidate_unwashed"
[5] "Candidate_washed"             "Control"
[7] "Standard_comparator_washed"
```

Now it is easier to see how the other treatment arms compare to the unwashed standard comparator. Superiority can be assessed using the p value for the fixed-effect parameter associated with the unwashed candidate net. Running the `summary()` function on the re-fitted model (here called `fit1a`, to distinguish from the original model) gives the following output (shortened for conciseness):

```
> fit1a <-
+   glm(
+     cbind(tot_dead, total - tot_dead) ~
+       treatment + hut + sleeper + day,
+     family = binomial, data = df)
> summary(fit1a)
Coefficients:
                                        Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept)                          -1.3001876  0.3456476  -3.762 0.000169 ***
treatmentActive_comparator_unwashed   1.2311743  0.1855108   6.637 3.21e-11 ***
treatmentActive_comparator_washed     0.6961056  0.1808125   3.850 0.000118 ***
treatmentCandidate_unwashed           1.2025083  0.1841820   6.529 6.62e-11 ***
treatmentCandidate_washed             0.6768528  0.1901508   3.560 0.000371 ***
treatmentControl                     -0.7850601  0.2210803  -3.551 0.000384 ***
treatmentStandard_comparator_washed -0.4415561  0.2280484  -1.936 0.052838 .
hut2                                 -0.3197885  0.1578667  -2.026 0.042797 *
hut3                                 -0.4449210  0.1603844  -2.774 0.005536 **
...
```

For superiority, we now need to check the p-value for the parameter for the unwashed candidate net:

```
> coef(summary(fit1a))[‘treatmentCandidate_unwashed’,"Pr(>|z|)"]
[1] 6.624834e-11
```

However, the alternative hypothesis that `glm()` uses is two-sided. Therefore, we should also check the **sign** of the coefficient for the unwashed candidate net: if it is positive then the candidate net is *superior* to the unwashed standard comparator, if it is negative then it is *inferior*.

```
> coef(summary(fit1a))[‘treatmentCandidate_unwashed’,"Estimate"]
[1] 1.202508
```

In this instance, we can conclude that the unwashed candidate is superior to the unwashed standard comparator, in terms of mosquito mortality.

The assessment for the washed candidate net is very similar to that carried out for the unwashed candidate. We shall not go through it here, but it is covered in the R code. The combined analysis is slightly different, so we will discuss it in detail. Let's look at the regression model we shall use here:

```
> fit3 <-
+   glm(
+     cbind(tot_dead, total - tot_dead) ~
+       ITN + hut + sleeper + wash + day,
+     family = binomial, data = df)
```

Now, the details of the trial arm (`treatment`) are provided by a combination of two variables: `ITN` and `wash`. As usual, before we run the model we should check which value of `ITN` will be used as the baseline case:

```
> levels(df$ITN)
[1] "Active_comparator"   "Candidate"       "Control"    "Standard_comparator"
```

This is fine for the non-inferiority assessment. Now run model `fit3` and look at the output (here truncated):

```
> summary(fit3)
...
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.075366   0.314884  -0.239 0.810838
```

```
ITNCandidate           -0.021888   0.111285  -0.197 0.844078
ITNControl             -2.003670   0.193250 -10.368  < 2e-16 ***
ITNStandard_comparator -1.189590   0.135565  -8.775  < 2e-16 ***
hut2                   -0.323268   0.157175  -2.057 0.039711 *
hut3                   -0.446850   0.160158  -2.790 0.005270 **
...
wash                   -0.513052   0.100605  -5.100 3.40e-07 ***
```

And, as before, we construct the OR and its 95%

```
OR3 <- exp(coef(summary(fit3))['ITNCandidate',"Estimate"])
OR3_lower <- exp(coef(summary(fit3))['ITNCandidate',"Estimate"] -
                   1.96*coef(summary(fit3))['ITNCandidate','Std. Error'])
OR3_upper <- exp(coef(summary(fit3))['ITNCandidate',"Estimate"] +
                   1.96*coef(summary(fit3))['ITNCandidate','Std. Error'])
```

And calculate the NIM (note that we use the mortality table for net type, not trial arm):

```
tab_mortality_ITN <- summm(df, vec = df$ITN, td = 'tot_dead',
                           tot = 'total', table = 1)
FIC_mortality3 <- tab_mortality_ITN[tab_mortality_ITN$Arm=='Active_comparator',
                        ]$Percentage / 100
non_inf_margin3 <- ((FIC_mortality3 - 0.07) / (1- (FIC_mortality3 - 0.07))) /
                        (FIC_mortality3 / (1- FIC_mortality3))
> non_inf_margin3
[1] 0.673509
```

Combining these things together, we find that the candidate net is non-inferior to the active comparator, for mosquito mortality (combining unwashed and washed nets together).

```
> plot_NI_OR(OR = OR3, ORl = OR3_lower, ORu = OR3_upper, mortality = 1,
+            NIM = non_inf_margin3, precision = 3)
[1] "OR=0.978 [0.787, 1.217]"
[1] "Non-inferior"
```

2.2.2 Blood Feeding

The procedure for assessing non-inferiority for blood feeding is extremely similar to that illustrated above for mosquito mortality. The only substantive difference stems from the fact that the effect of the vector control products is to reduce blood feeding, compared to increasing mosquito mortality. Reducing blood feeding is sometimes described as increasing blood-feeding inhibition. We will try to keep our descriptions clear in what follows.

To illustrate the process, we will go through Assessment 4 (comparing the unwashed candidate net to the unwashed active comparator). Here we write down the regression model to be used, first checking that the model has the correct baseline treatment arm:

```
df$treatment <- relevel(df$treatment, 'Active_comparator_unwashed')
levels(df$treatment)


fit4 <-
```

```
  glm(
    cbind(tot_bf, total - tot_bf) ~
      treatment + hut + sleeper + day,
    family = binomial, data = df)
summary(fit4)
```

Notice that now we are looking at the proportion of collected mosquitoes that have blood fed. Let's take a look at some of the model output:

```
summary(fit4)
                                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                         -1.121312   0.284960  -3.935 8.32e-05 ***
treatmentStandard_comparator_unwashed  0.084849   0.180408   0.470 0.638128
treatmentActive_comparator_washed      0.587605   0.164835   3.565 0.000364 ***
treatmentCandidate_unwashed           -0.009009   0.182133  -0.049 0.960548
treatmentCandidate_washed              0.643385   0.172984   3.719 0.000200 ***
treatmentControl                       0.956698   0.160061   5.977 2.27e-09 ***
treatmentStandard_comparator_washed    0.473101   0.175575   2.695 0.007048 **
hut2                                  -0.183247   0.153882  -1.191 0.233723
hut3                                  -0.259986   0.155708  -1.670 0.094978 .
...
```

We calculate the odds ratio and 95% CIs in exactly the same way as for mosquito mortality:

```
OR4 <- exp(coef(summary(fit4))['treatmentCandidate_unwashed',"Estimate"])
OR4_lower <- exp(coef(summary(fit4))['treatmentCandidate_unwashed',"Estimate"] -
            1.96*coef(summary(fit4))['treatmentCandidate_unwashed','Std. Error'])
OR4_upper <- exp(coef(summary(fit4))['treatmentCandidate_unwashed',"Estimate"] +
            1.96*coef(summary(fit4))['treatmentCandidate_unwashed','Std. Error'])
```

We now use the proportion of mosquitoes blood fed in the unwashed active comparator arm to set the NIM. Recall that we use the estimate taken directly from the dataset for this, rather than the regression model:

```
FIC_bf4 <- tab_bf[tab_bf$Arm=='Active_comparator_unwashed',]$Percentage / 100
non_inf_margin4 <-
        ((FIC_bf4 + 0.07) / (1- (FIC_bf4 + 0.07))) / (FIC_bf4 / (1- FIC_bf4))
```

Note that now the OR for the NIM is selected based on blood feeding being no more than 7% higher for the unwashed candidate net. For non-inferiority, the entire 95% CI for the OR must lie *below* the NIM. The user can assess this either by inspection of the values calculated above, or by using the `plot_NI_OR()` function. We set the argument `mortality` equal to 0, to indicate that this assessment is for blood-feeding inhibition:

```
plot_NI_OR(OR = OR4, ORl = OR4_lower, ORu = OR4_upper, mortality = 0,
           NIM = non_inf_margin4, precision = 3)
[1] "OR=0.991 [0.694, 1.416]"
[1] "Non-inferior"
```

Hence we have demonstrated non-inferiority of the unwashed candidate net compared to the unwashed active comparator, for blood feeding.

We should also check that the unwashed candidate net is superior to the unwashed standard comparator:

```
#Change baseline treatment arm
df$treatment <- relevel(df$treatment, 'Standard_comparator_unwashed')
levels(df$treatment)
[1] "Standard_comparator_unwashed" "Active_comparator_unwashed"
[3] "Active_comparator_washed"     "Candidate_unwashed"
[5] "Candidate_washed"     "Control"
[7] "Standard_comparator_washed"
fit4a <-
  glm(
    cbind(tot_dead, total - tot_dead) ~
      treatment + hut + sleeper + day,
    family = binomial, data = df)
summary(fit4a)
#Check the p-value
coef(summary(fit4a))['treatmentCandidate_unwashed',"Pr(>|z|)"]
```

Here we see that the relevant p-value is much less than 0.05. Finally we should check that the coefficient is positive (*i.e.* the candidate net is superior, not inferior):

```
coef(summary(fit4a))['treatmentCandidate_unwashed',"Estimate"] < 0
[1] TRUE
```

Although we have not gone through all 6 non-inferiority assessments listed at the start of this Section, the examples given should allow the user to generate the remaining assessments (they are fully specified in the R code which accompanies this tutorial). Table 1 summarises the assessments, for completeness.

2.3 Summarising and plotting the data

Often when carrying out a non-inferiority (or superiority) assessment, it is useful to also provide a summary of the trial data (total number of mosquitoes collected in each arm, model-adjusted mosquito mortality, *etc.*). We have developed some functions to tabulate and visualise the data. However, there are some subtleties around interpreting the output of the regression models, which we will discuss.

Let's first look at the mosquito mortality estimated directly from the dataset. For each trial arm, this is simply the total number of dead mosquitoes divided by the total number of mosquitoes. For this dataset, we find:

```
> summm(df, vec = df$treatment, td = 'tot_dead', tot = 'total')
[1] "Control: 5.99%"
[1] "Standard_comparator_unwashed: 11.61%"
[1] "Standard_comparator_washed: 8.75%"
[1] "Active_comparator_unwashed: 34.09%"
[1] "Active_comparator_washed: 21.45%"
[1] "Candidate_unwashed: 28.99%"
[1] "Candidate_washed: 22.54%"
```

Now let's compare this to the estimates generated from the regression model. The function `XX()` has been designed to do this for all trial arms (see appendix for full details):

```
> mFE(model = fit1, vec = df$treatment, intercept = 'Active_comparator_unwashed',
    bfi = 0, name = 'treatment')
                          Arm Mortality Lower_95pc_CI Upper_95pc_CI
2                     Control     0.111         0.059         0.197
1   Active comparator unwashed     0.483         0.333         0.635
3 Standard comparator unwashed     0.214         0.127         0.338
4   Standard comparator washed     0.149         0.084         0.250
5     Active comparator washed     0.353         0.223         0.510
6           Candidate unwashed     0.476         0.316         0.640
7             Candidate washed     0.349         0.214         0.514
```

These results are quite different to the data-derived estimates. This is because we have not considered the role of the other fixed effects- day, hut and sleeper. To be precise, the estimates above are the model-estimated mortalities in each treatment arm in hut 1, day 1 and with sleeper 1. If the observed mortalities for this combination of fixed effects is not typical of that observed across the whole trial, then the summarised mortalities may look at bit strange. We could instead present all the model parameters, which would provide a comprehensive overview of the fitted model. However, in this case we have 6 parameters for hut, 6 parameters for sleeper and 48 parameters for day. So it is a lot of information to present- and a lot of information for the reader to absorb. Let's think a bit more carefully about how we set up the model, and how we choose the baseline categories for the fixed effects in our model. From a logical point of view, it should make no difference which categories are chosen as the baseline categories for the model. As we have 7 huts, 7 sleepers and 49 trial days, there are $7 \times 7 \times 49 = 2401$ ways of setting up the baseline category for the model. We have developed a bespoke function that looks at all these permutations for the baseline category for the model and calculates the estimates (on the log-odds scale) for the mosquito mortality for the baseline category, across all possible 2401 estimates. The function returns an 'offset', which adjusts the model output to return the *median* value for the mosquito mortality. We demonstrate the procedure as follows:

```
> ofs1 <- new_median_FE(model = fit1, FE = c(‘hut’,‘sleeper’,‘day’))
> ofs1
[1] -0.9085774
```

The argument `FE` must list all the fixed effects we wish to consider. We now re-run the function `mFE()`, with this offset as one of the arguments:

```
> mFE(model = fit1, vec = df$treatment, intercept = ‘Active_comparator_unwashed’,
            bfi = 0, name = "treatment", offset = ofs1)
                          Arm Mortality Lower_95pc_CI Upper_95pc_CI
2                     Control     0.048         0.025         0.090
1   Active comparator unwashed     0.273         0.168         0.413
3 Standard comparator unwashed     0.099         0.055         0.170
4   Standard comparator washed     0.066         0.036         0.118
5     Active comparator washed     0.181         0.104         0.295
6           Candidate unwashed     0.268         0.157         0.417
7             Candidate washed     0.178         0.099         0.299
```

These values are closer to the values directly from the dataset. The output from this function could be used to generate a table e.g.

```
tbl <- mFE(model = fit1, vec = df$treatment,
        intercept = 'Active_comparator_unwashed',
        bfi = 0, name = "treatment", offset = ofs1)
#save as a .csv file
write.csv('Table_mosquito_mortality.csv',tbl)
```

Now let's prepare the data for a visualisation. If we only wish to show the data from unwashed ITNs, we can remove the washed arms:

```
mk1a <- mk1[-grep(" washed", mk1$Arm),]
#Determine the plotting order
mk1a$ord <- c(1,3,2,4) #Should match the order of the labels in the legend
p1 <- ggplot(data = mk1a) +
  geom_errorbarh(aes(y = ord, xmin = Lower_95pc_CI,
    xmax = Upper_95pc_CI), height = 0) +
  geom_point(aes(y=ord, x=Mortality, colour = Arm), size = 3) +
  xlim(c(0,1)) + xlab('Proportion of mosquitoes blood fed') +
  theme_classic() + ylab('') +
  theme(axis.line.y = element_blank(),
    axis.ticks.y = element_blank(), axis.text.y = element_blank()) +
  scale_color_discrete(breaks = c('Candidate unwashed',
    'Active comparator unwashed', 'Standard comparator unwashed','Control')) +
  theme(legend.position = c(0.8,0.3)) + labs(color = '') +
  ggtitle('Mosquito mortality (unwashed ITNs)')
```

Here we've saved the plot as `p1`. If we save the non-inferiority assessment as `NI_1`, we can display the plots together:

```
NI_1 <- plot_NI_OR(OR = OR1, ORl = OR1_lower, ORu = OR1_upper, mortality = 1,
        NIM = non_inf_margin1, precision = 3)
plot_grid(p1,NI_1,nrow = 1, rel_widths = c(0.6,0.4), labels = c('A','B'))
```

These plots are shown in Figure 2.2.

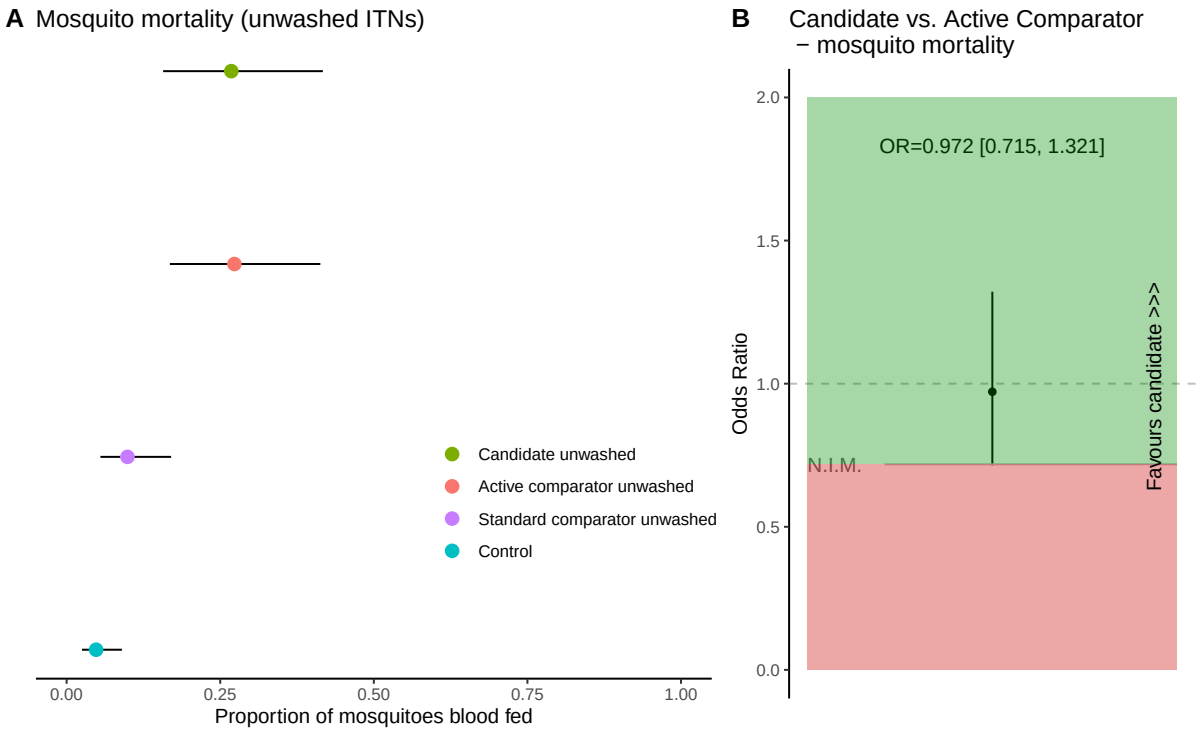


Figure 2.2: A visualisation of the non-inferiority assessment: comparing the unwashed candidate net to the unwashed active comparator. Panel A: mortality estimates (and 95% confidence intervals) in the arms containing unwashed nets, as well as the untreated control. Panel B: A summary of the non-inferiority assessment between the unwashed candidate net and the unwashed active comparator (for mosquito mortality). In this instance, the non-inferiority margin (N.I.M., horizontal grey line) is very close to the 95% CI: therefore, the numerical values of these quantities must be inspected, to make the non-inferiority assessment (see Table 2.1).

	Assessment	N.I.M.	Odds Ratio (95% CI)	Superior to the standard comparator?
1.	Mortality: unwashed candidate *vs.* unwashed active comparator	0.718	0.972 [0.715, 1.321]	Yes
2.	Mortality: washed candidate *vs.* washed active comparator	0.619	0.981 [0.717, 1.341]	Yes
3.	Mortality: candidate *vs.* active comparator (combined)	0.674	0.978 [0.787, 1.217]	Yes
4.	Blood feeding: unwashed candidate *vs.* unwashed active comparator	1.541	0.991 [0.694, 1.416]	No
5.	Blood feeding: washed candidate *vs.* washed active comparator	1.404	1.057 [0.795, 1.406]	No
6.	Blood feeding: candidate *vs.* active comparator (combined)	1.446	1.024 [0.821, 1.278]	No

Table 2.1: A summary of the six non-inferiority assessments carried out on the dataset used for this tutorial. We show the non-inferiority margin (N.I.M.) used in each case. For mosquito mortality, the entire 95% confidence interval for the odds ratio must lie above the N.I.M. for non-inferiority to be shown. For blood-feeding, the entire 95% confidence interval must lie below the N.I.M. for non-inferiority to be shown. We also check that the candidate net is superior to the standard comparator (rightmost column).

Chapter 3

Analysis in STATA

3.1 Loading & summarising the dataset

In this chapter, we will repeat the same non-inferiority assessments that we carried out in R, for the same dataset but this time in STATA[1]. We first load the data (saved as a `.csv` file) using the `import` command:

```
import delimited example_dataset.csv
```

Note that you may need to change the current working directory before you do this, *e.g.*

```
cd C:\Users\username\Documents
```

Here the file path should match the location of the materials for this tutorial. You can run the command `pwd` if you need to check the current working directory. Once the data is loaded, we can take a look at it. The command `tab` gives the frequency counts for a numerical variable:

```
. tab tot_dead
```

tot_dead	Freq.	Percent	Cum.
0	117	34.11	34.11
1	81	23.62	57.73
2	48	13.99	71.72
3	32	9.33	81.05
4	17	4.96	86.01
5	16	4.66	90.67
6	10	2.92	93.59
7	7	2.04	95.63
8	3	0.87	96.50
9	4	1.17	97.67
10	3	0.87	98.54
11	2	0.58	99.13
14	1	0.29	99.42

[1]These analysis have been run in both STATA version 13 and STATA version 19.

```
        15 |          2        0.58      100.00
-----------+-----------------------------------
     Total |        343      100.00
```

Whereas the command `summarize` gives the summary statistics for the variable:

```
. summarize tot_dead

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    tot_dead |       343           2    2.575185          0         15
```

The command `levelsof` returns a list of the values of a variable, textite.g.

```
. levelsof(treatment)
`"Active_comparator_unwashed"' `"Active_comparator_washed"'
`"Candidate_unwashed"' `"Candidate_washed"' `"Control"'
`"Standard_comparator_unwashed"' `"Standard_comparator_washed"
```

3.2 Making the non-inferiority assessment

3.2.1 Mosquito mortality

Now we will reproduce the non-inferiority assessments we carried out in R in the previous chapter, and check that we can obtain the same results. We will start by comparing the unwashed candidate to the unwashed active comparator in terms of mosquito mortality (this was Assessment 1 in the previous chapter). Our first step will be to calculate the NIM. We use the `collapse` function to sum together the mosquito counts (both total mosquitoes and dead mosquitoes) for the various treatment arms. We then calculate the (unadjusted) mosquito mortality observed in each arm:

```
collapse (sum) sum1=tot_dead sum2=total, by(treatment)
gen prop_dead = sum1/sum2
list
```

```
     +-------------------------------------------------------+
     |                    treatment   sum1   sum2   prop_d~d |
     |-------------------------------------------------------|
  1. |   Active_comparator_unwashed    150    440   .3409091 |
  2. |     Active_comparator_washed    133    620   .2145161 |
  3. |           Candidate_unwashed    147    507   .2899408 |
  4. |             Candidate_washed    112    497   .2253521 |
  5. |                     Control     42    701   .0599144 |
     |-------------------------------------------------------|
  6. | Standard_comparator_unwashed     60    517   .1160542 |
  7. |   Standard_comparator_washed     42    480      .0875 |
     +-------------------------------------------------------+
```

Recall that the NIM here should be set so that the proportion of mosquitoes killed by the unwashed candidate net should be no more than 7% less than the proportion killed by the unwashed active comparator. We construct the NIM as follows:

```
gen or1 = (prop_dead - 0.07)/(1-prop_dead + 0.07)
gen or2 = (prop_dead)/(1-prop_dead)
*Calculate the odds-ratio (OR) for the NIM
gen nim = or1/or2
list
```

treatment	sum1	sum2	prop_d~d	or1	or2	nim
Active_co~unwashed	150	440	.3409091	.3715711	.5172414	.7183707
Active_co~washed	133	620	.2145161	.1689291	.2731006	.61856
Candidate_unwashed	147	507	.2899408	.2819541	.4083333	.6904997
Candidate_washed	112	497	.2253521	.1839253	.2909091	.6322432
Control	42	701	.0599144	-.0099849	.0637329	-.1566677
Standard_~unwashed	60	517	.1160542	.0482775	.131291	.3677139
Standard_~washed	42	480	.0875	.0178117	.0958904	.1857506

Note that, in this case, it is the NIM for the unwashed active comparator that we need to extract from this table (`0.718307...`). We will now reload the original dataset, to fit the regression model. We will first save the information generated above, and append it to the full dataset:

```
% save in the current working directory
save "aggregated_mortality.dta"
clear
import delimited example_dataset.csv
append using aggregated_mortality.dta
% we can drop the variables that we won't use again
drop sum1 sum2 prop_dead or1 or2
```

Note that we will also need to make an equivalent table for blood-feeding (this can be found in the STATA code that accompanies this tutorial). We are nearly ready to fit the regression model. However, STATA has imported `treatment` as a 'string' variable. We will need to generate an equivalent 'factor' variable to include in the regression model. We will call this `treatment2`:

```
encode(treatment), generate(treatment2)
\*We only want treatment2 values for the original dataset
(not the values we've appended beneath):*/
replace treatment2=. if day==.
```

Here we can see how the levels of treatment2 correspond to treatment:

```
. label list treatment2
treatment2:
           1 Active_comparator_unwashed
           2 Active_comparator_washed
           3 Candidate_unwashed
           4 Candidate_washed
           5 Control
           6 Standard_comparator_unwashed
           7 Standard_comparator_washed
```

We will use the `blogit` command to fit the regression model, as this fits a logistic regression model to aggregated count data. Including fixed-effects for treatment, hut, sleeper, and day, the model has the following form:

```
blogit tot_dead total i.treatment2 i.hut i.sleeper i.day
```

Remember that for these models it is important to consider which `treatment2` category is used as the baseline. By default, category 1 is used, which in this case is the best one for us to use. We will show an example later, where we will manually change the baseline category. Once the model is run, the results will be stored in STATA's memory. We won't explore the whole model contents here, but it can be viewed by running the command `ereturn list`. Now we will construct the ORs for the non-inferiority assessment. We need to extract the parameter estimate for the unwashed candidate net category for `treatment`. This is category 3 of `treatment2`. STATA stores the estimates for the fixed effects in the container `_b[]`, whilst the standard errors for these parameters are stored in `_se[]`. We extract the OR and its 95% CI like this:

```
gen or_model = exp(_b[_outcome:3.treatment2])
gen or_model_lower = exp(_b[_outcome:3.treatment2] - 1.96* _se[_outcome:3.treatment2])
gen or_model_upper = exp(_b[_outcome:3.treatment2] + 1.96* _se[_outcome:3.treatment2])
```

The `display` command will print these values to the screen:

```
. display or_model
.97174102
. display or_model_lower
.71467191
. display or_model_upper
1.3212785
```

An alternative way to extract the OR is to request the ORs when fitting the model (output truncated):

```
. blogit tot_dead total i.treatment2 i.hut i.sleeper i.day, or

------------------------------------------------------------------------------
         _outcome | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------------------+-----------------------------------------------
       treatment2 |
  Active~_washed  |   .5856291   .0904798    -3.46   0.001     .4326251     .792745
Candid~_unwashed  |    .971741   .1523381    -0.18   0.855      .714676    1.321271
Candidate_washed  |   .5744619   .0950874    -3.35   0.001     .4153045    .7946132
         Control  |   .1331559   .0269214    -9.97   0.000     .0895912    .1979046
Standard~unwashed |   .2919496   .0541598    -6.64   0.000     .2029555    .4199667
  Standard~washed |   .1877338   .0382263    -8.21   0.000     .1259566    .2798104
                  |
              hut |
               2  |   .7263026    .114659    -2.03   0.043     .5330172    .9896783
               3  |   .6408749   .1027864    -2.77   0.006     .4680086    .8775922
...
```

Regardless of which method we use, we can see that the 95% CI `[.714676-1.321271]` is not entirely above the NIM we calculated above (`0.718307...`). Therefore, we cannot conclude that the unwashed candidate is non-inferior to the unwashed active comparator, in terms of mosquito mortality. This is consistent with what we found in R (Table 2.1). For completeness, however, we will check whether the unwashed candidate is superior to the unwashed standard comparator. We can use the same regression model again: but it will be simpler if we change the `treatment2` baseline category to be the unwashed standard comparator. This is level 6 of `treatment2` (output truncated):

```
. blogit tot_dead total ib6.treatment2 i.hut i.sleeper i.day
------------------------------------------------------------------------------
         _outcome |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------------+---------------------------------------------------
       treatment2 |
Active~_unwashed  |  1.231174   .1855108     6.64   0.000     .8675798    1.594769
Active_c~_washed  |  .6961056   .1808125     3.85   0.000     .3417195    1.050492
Candidate_unwashed|  1.202508    .184182     6.53   0.000     .8415183    1.563498
 Candidate_washed |  .6768528   .1901508     3.56   0.000     .3041641    1.049541
         Control  | -.7850601   .2210804    -3.55   0.000     -1.21837   -.3517505
Standard_~_washed | -.4415561   .2280484    -1.94   0.053    -.8885226    .0054105
                  |
              hut |
                2 | -.3197885   .1578667    -2.03   0.043    -.6292016   -.0103754
                3 |  -.444921   .1603844    -2.77   0.006    -.7592687   -.1305733
```

Here we see that the coefficient for the unwashed candidate is positive, and that the p-value is <0.001. So we can conclude that the unwashed candidate net is superior to the unwashed standard comparator, in terms of mosquito mortality. We won't generate the arm-level mortality estimates here, but we will show an example of how to generate these from the fitted regression model. STATA denotes the intercept for the regression model with the label `_cons`. We can calculate the mosquito mortality for the intercept category in the most-recently fitted regression model using the `invlogit` command:

```
. display invlogit(_b[_cons])
.21413345
```

So this means that the estimated mosquito mortality in the unwashed standard comparator arm, on day 1 in hut 1 with sleeper 1 is about 21.4%.

3.2.2 Blood feeding

The procedure for assessing non-inferiority for mosquito blood feeding is extremely similar to that for mosquito mortality. However, we shall go through the process here, for completeness. As we assessed unwashed nets for mosquito mortality, we will perform the combined analysis here (this is Assessment 6 in the code). As before, our first step is to calculate the NIM. Note that we now group the data by the variable `itn`:

```
clear
import delimited "example_dataset.csv"
collapse (sum) sum1=tot_bf sum2=total, by(itn)
```

```
gen prop_fed = sum1/sum2
/*Note: NIM chosen by considering slightly higher blood feeding for the
candidate net (compared to slightly lower mosquito mortality)*/
gen or1 = (prop_fed + 0.07)/(1-prop_fed - 0.07)
gen or2 = (prop_fed)/(1-prop_fed)
*Calculate the odds-ratio (OR) for the NIM
gen nim = or1/or2
list
save "aggregated_bf_itn.dta"
```

Now we reload the data, and append the NIM, as before:

```
clear
import delimited "example_dataset.csv"
append using "aggregated_bf_itn.dta"
*Remove variables we don't need anymore
drop sum1 sum2 prop_fed or1 or2
```

Now we generate a new factor variable, `itn2`, to use in the regression model:

```
encode(itn), generate(itn2)
replace itn2=. if day==.
```

For the regression model, we again use the command `blogit`. This time we use `tot_bf` in the model, which is the number of mosquitoes that were blood fed:

```
blogit tot_bf total i.itn2 i.hut i.sleeper i.day i.wash
```

Note that we have included an additional fixed effect for the washed status of the net. Now we calculate the OR and its 95% confidence intervals

```
gen or_model = exp(_b[_outcome:2.itn2])
gen or_model_lower = exp(_b[_outcome:2.itn2] - 1.96* _se[_outcome:2.itn2])
gen or_model_upper = exp(_b[_outcome:2.itn2] + 1.96* _se[_outcome:2.itn2])
```

Finally, we extract the relevant NIM from the dataset with the following command:

```
list if itn=="Active_comparator" & missing(day)
```

This yields an NIM of 1.44578. Comparing it to the 95% CI for the OR:

```
. display or_model_lower
.82099819
. display or_model_upper
1.2784339
```

we can see that we have demonstrated non-inferiority for blood feeding. We should also check for superiority to the standard comparator. We can do this with the same regression model, changing the baseline category for `itn2` (output truncated):

```
. blogit tot_bf total ib4.itn2 i.hut i.sleeper i.day i.wash
------------------------------------------------------------------------------
     _outcome |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
         itn2 |
Active_com~r  |   .0264738   .1149307     0.23   0.818    -.1987863    .2517339
   Candidate  |   .0506745   .1182441     0.43   0.668    -.1810796    .2824287
     Control  |   .9533925   .1320353     7.22   0.000     .6946081    1.212177
              |
          hut |
            2 |  -.1709663   .1532829    -1.12   0.265    -.4713952    .1294627
```

Here we find that the candidate is not superior to the standard comparator for blood feeding. This matches our findings in the R analysis (Table 2.1).

Appendix: A description of the R functions developed for this work

- `plot_NI_OR()`
  - Description:
  - Usage: `plot_NI_OR(OR, ORl, ORu, mortality, NIM, precision, title)`
  - Arguments:
    - `OR, ORl, ORu`: These are: the odds ratio estimate, and the lower and upper estimates of its 95%, respectively
    - `mortaltiy`: Set equal to 1 if the non-inferiority assessment is for mosquito mortality; set to zero for blood feeding. The default option is 1.
    - `NIM`: The non-inferiority margin. This should be in the form of an odds ratio. For blood-feeding, a warning will be issued if `NIM<1`. Similarly, for mosquito mortality a warning will be issued if `NIM>1`
    - `precision`: the number of decimal places to be used for the non-inferiority summary that is printed on the visualisation. Note: unrounded values are used for the odds-ratios in the actual non-inferiority assessment.
    - The title to be used for the visualisation
- `summm()`
  - Description: A simple function to summarise the mosquito mortality or blood feeding observed in each trial arm
  - Usage: `summm(data, vec, td, tot, table, precision)`
  - Arguments:
    - `data`: Dataset to be analysed
    - `vec`: Variable in the dataset we wish to use to stratify the data. If this variable is `treatment` and the dataset is `df`, we write: `vec=df$treatment`
    - `td`: This is the numerator for the summary. For mosquito mortality, this is `tot_dead`; for blood-feeding it is: `tot_bf`
    - `tot`: This is the denominator for the summary. Here this will be `total`
    - `table`: This variable determines the form of the output. For `table=0`, sentences will be printed in the R console; otherwise the function returns a data frame.
    - `precision`: number of dps (sig figs?) in the output
- `mFE()`
  - Description: A function that takes a fitted regression model and returns the mortality (blood feeding) estimates, along with their 95% CIs.
  - Usage: `mFE(model, vec, intercept, bfi, name, offset)`
  - Arguments:
    - `model`: Fitted regression model
    - `vec`: Variable in the dataset we wish to use to stratify the data. If this variable is `treatment` and the dataset is `df`, we write: `vec=df$treatment`
    - `intercept`: In the regression model output, the treatment arm that pertains to the baseline category is not named. So we must enter it here.

    * `bfi`: If `bfi=0`, the function will assume the model is estimating mosquito mortality. For `bfi=1`, it will assume the model is estimating blood feeding. The default value is 0.
    * `name`: Should match the variable entered in the argument `vec`. So if `vec=df$treatment`, we write `name = 'treatment'`
    * `offset`: Adjustment factor (on the log-odds scale) to the mortality (or blood-feeding) measured in the baseline category (Default value is 0).

- `new_median_FE()`
  - Description: A function which looks at the permutations of fixed-effects that gives a representative mortality (or blood feeding) estimate for the baseline category. This is done by calculating the median value, across all permutations
  - Usage: `new_median_FE(model, FE)`
  - Arguments
    * `model`: Fitted regression model
    * `FE`: At the moment, this argument can only take lists of containing 1,2, or 3 fixed effects. The default list is `FE = c('hut','sleeper','day')`

Malaria Policy Advisory Group Meeting
30 October–1 November 2023, Geneva, Switzerland
Background document for Session 4

Revised pathways for the evaluation process of vector control products for malaria

This draft flowchart illustrates the processes by which the World Health Organization (WHO) evaluates vector control products for malaria, with a focus on illustrating the use of comparative efficacy data in this context. The "prequalification" and "new intervention" pathways remain largely unchanged from previous communications; however, the latter will start to include an application type for investigational new vector control products (INVCs) to ensure the quality, safety and efficacy of product batches used in epidemiological trials.

Comparative efficacy data will be required for all products other than those evaluated in epidemiological trials (and excluding pyrethroid-only ITNs, which are in the process of being replaced by more effective nets). The aim of non-inferiority assessments conducted with these data is to validate the applicability of existing recommendations or to identify and inform the need to extend an existing recommendation or develop a new one. Which of these outcomes apply will be determined by the Guideline Development Group, with support from a guidelines methodologist, using the process described in the *WHO handbook for guideline development, second edition*. The information presented here has been developed by the Vector Control Unit of the WHO Global Malaria Programme, in consultation with the WHO Guidelines Review Committee and the Prequalification Team for Vector Control Products.

WHO teams: Technical department | PQT | GRC | PCC

WHO-convened groups: VCAG | GDG

External groups: Applicant

Mandatory route

Route dependent on relevant status

Responsible team: action point

Outcome of respective action

Applicants: Request for determination of pathway for new intervention submitted to WHO

Pre-submission coordination committee: Assign intervention to intervention class, and appropriate pathway based on data requirements. Communicate outcome to applicant.

PQT: requirement of safety, quality and efficacy of intervention

ESTABLISHED INTERVENTION CLASS for which directly related or similar WHO recommendation(s) exist

TENTATIVE INTERVENTION CLASS for which no WHO recommendation currently exists

Prequalification

PQT: presubmission guidance on dossier requirements

Applicant: Prepare dossier according to requirements, relating to safety, quality, entomological efficacy (*and non-inferiority if directed*)

PQT: review submitted safety and quality data of product

PQT: review submitted efficacy data of product

PQT: internal review

Prequalification decision: Meets requirements

Prequalification decision: Assessment failure

Applicant: encouraged to submit new dossier addressing deficiencies assessed in first assessment

PQT: prepares WHO public assessment report

Comparative efficacy

SIC is very similar to FIC product in all characteristics

SIC is has either different active ingredients or different use patterns compared to FIC

SIC product that combines elements/active ingredients covered by multiple existing recommendations

Technical department: inform applicant of appropriate active comparator for testing non-inferiority

Technical department: inform applicant of appropriate active comparator for testing non-inferiority. Submit planning proposal to GRC for the relevant intervention class(es) that may be updated.

Applicant: Non-inferiority studies conducted to compare SIC to assigned FIC active comparator product from the respective intervention class

Applicant: Non-inferiority studies conducted to compare SIC to assigned active comparator product(s)

Applicant: Non-inferiority studies conducted to compare SIC to assigned FIC active comparators from each established intervention class

Technical department: review data submitted on non-inferiority

Non-inferiority demonstrated via ≥ 2 well conducted studies

Non-inferiority not demonstrated

Non-inferiority demonstrated via ≥ 2 well conducted studies

Non-inferiority not demonstrated

Non-inferiority demonstrated via ≥ 2 well conducted studies

Non-inferiority not demonstrated

Applicant: GMP to hold further discussion with the applicant on next steps

Applicant: GMP to hold further discussion with the applicant on next steps

Applicant: GMP to hold further discussion with the applicant on next steps

GDG: draw on NI data & contextual factors for EtD discussion

GDG: draw on NI data & contextual factors for EtD discussion

GDG: develops new or extends current recommendation

GDG develops new recommendation

New intervention

Novel intervention (eventually may become FIC product)

Technical department: submit planning proposal to GRC for new recommendation for intervention class

Applicant: submission of study documents to VCAG in preparation for trials

Applicant: Compile INVC package to ensure intervention used in trials is safe and consistent for experimental use.

VCAG: review trial protocols and study documents, provide applicants with scientific advice on generation of evidence of disease impact

PQT: review data submitted for safety of intervention for experimental use

Applicant: Conduct epidemiological trials (at least 2 needed for each intervention class). Trials submitted to VCAG for assessment

Applicant: submission of manufacturing batch data to validate intervention used in trials

VCAG: Public health value assessed via ≥ 2 trials

Public health value not assessable due to trial limitations/ statistical power

Public health value not demonstrated, despite trials being of high quality

Public health value demonstrated via ≥ 2 well conducted studies

GDG: review publicly available data & contextual factors relevant to the intervention class and evaluate EtD discussion

GDG develops new recommendation

GRC: reviews and approves proposed guidelines revision

GRC: reviews and approves proposed new guidelines

Technical department: prepares guideline update

RECOMMENDATION and PREQUALIFICATION issued

Comparative effectiveness in the context of the arrival of new vector control products

Malaria Policy Advisory Group
Geneva, Switzerland

31 October 2023

Dr Jan Kolaczinski, Unit Head, Vector Control & Insecticide Resistance

World Health Organization

Content

- Technical consultation
- Guidelines update
- Protocol update
- Evaluation & guidelines process update
- Next steps

World Health Organization

April 2023

VECTRON T500

Product Identification

Product Type:	IRS
PQT/VC Ref Number:	P-03226
Applicant:	
Other names:	
Active Ingredient/Synergist:	
Concentration:	
Formulation Type:	
Supporting WHO Recommendations:	

Prequalification Status

Status of Prequalification:	
Date of Prequalification:	
Basis of Listing:	

PermaNet Dual

Product Identification

Product Type:	ITN
PQT/VC Ref Number:	P-03228

in; 5.0 g/kg (200 mg/m2) Chlorfenapyr

World Health Organization

Health Topics | Countries | Newsroom | Emergencies | Data | About WHO

Home / Newsroom / Article / Entomological data to allow comparative assessment of insecticide-treated

Entomological data to allow comparative assessment of insecticide-treated nets and indoor residual spraying products

Deadline for submission: 19 May 2023

17 March 2023 | Call for data

Comparative assessments in the context of the arrival of new vector control products: MPAG welcomed the update from the Global Malaria Programme on the data requirements to support development and implementation of normative guidance for new vector control products. MPAG reiterated its earlier guidance, first issued in 2017, that comparative assessments of entomological data are required for all products other than the "first-in-class" products that generate the epidemiological data used to establish an intervention class. Given the need to balance rapid market access to new products with the need for rigorous data that demonstrate comparative effectiveness relative to existing products, MPAG requested that the Global Malaria Programme urgently clarify and resolve issues associated with the implementation of this process. MPAG emphasized the need to ensure that there is a single coordinated process for WHO to evaluate and approve new products. Once internal implementation issues have been resolved between the Global Malaria Programme and the Prequalification Team, MPAG considers it important for the process to be better communicated to external stakeholders to ensure consistent messaging and a common understanding of the data required for new products.

https://www.who.int/publications/i/item/9789240074385

World Health Organization

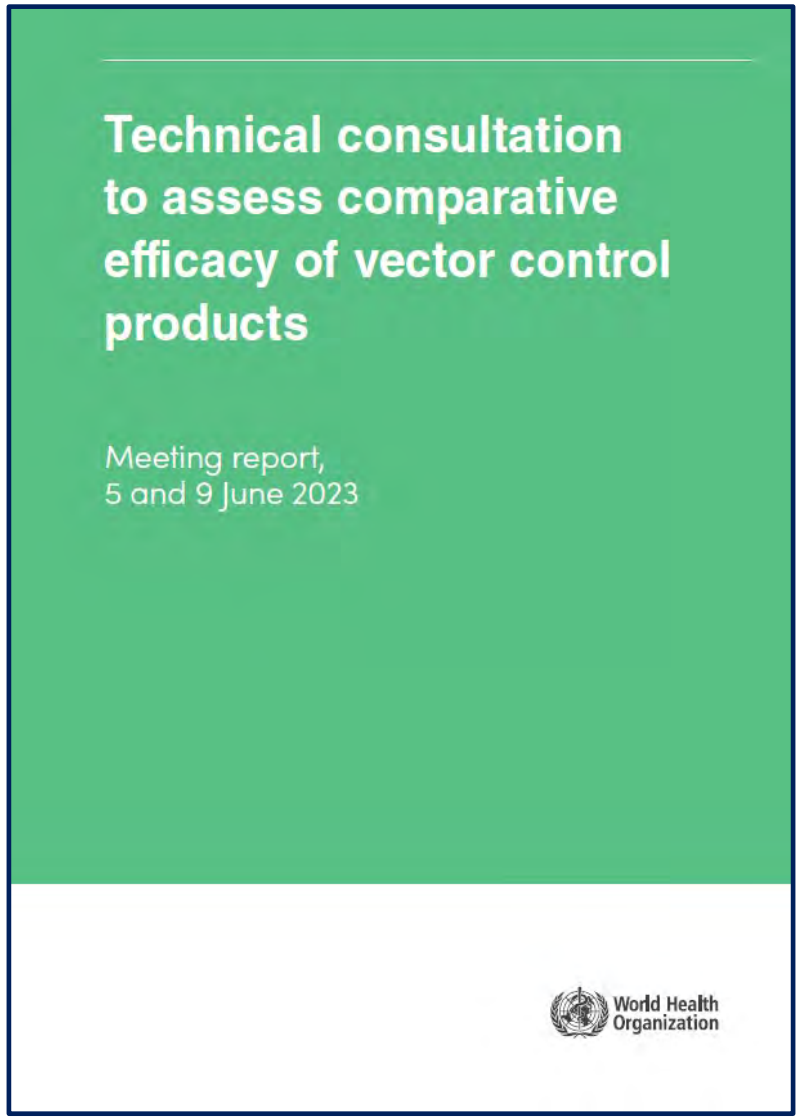
Technical consultation on:

DuraNet Plus©
(alpha-cypermethrin+PBO) compared to Olyset™ Plus (permethrin+PBO)

Yorkool® G3
(deltamethrin+PBO) compared to Olyset™ Plus

PermaNet® Dual
(deltamethrin+chlorfenapyr) compared to Interceptor® G2 (alpha-cypermethrin+chlorfenapyr)

VECTRON™ T500
(Meta-diamides: Broflanilide) compared to Actellic 300CS (organophosphate: pirimiphos-methyl)



https://www.who.int/publications/i/item/9789240078659

World Health Organization

Technical consultation outputs:

Product	Findings	Recommendation to WHO
DuraNet Plus©	Non-inferior to Olyset™ Plus for mosquito mortality based on pooled (washed and unwashed) data	Consider DuraNet Plus© and Yorkool® G3 to be covered under the current pyrethroid-PBO net recommendation
Yorkool® G3	Non-inferior to Olyset™ Plus for mosquito mortality based on pooled (washed and unwashed) data	
PermaNet® Dual	Non-inferior to Interceptor® G2 for mosquito mortality based on pooled (washed and unwashed) data	Consider PermaNet® Dual to be covered under the current pyrethroid-chlorfenapyr net recommendation
VECTRON™ T500	Not non-inferior to Actellic 300CS for mosquito mortality at three months, but non-inferior after at six months time point	Extend IRS recommendation to include broflanilide, in turn covering VECTRON™ T500 under this recommendation and making it the appropriate active comparator for other broflanilide products in future comparative efficacy assessments

World Health Organization

Guidelines Update

Insecticide formulations currently recommended by WHO for use in IRS:

Sodium channel modulators

Pyrethroids: alphacypermethrin, deltamethrin, lambda-cyhalothrin, etofenprox, bifenthrin

Organochlorines (e.g. DDT): no prequalified products available

Acetylcholinesterase inhibitors

Organophosphates: pirimiphos-methyl

Carbamates: bendiocarb

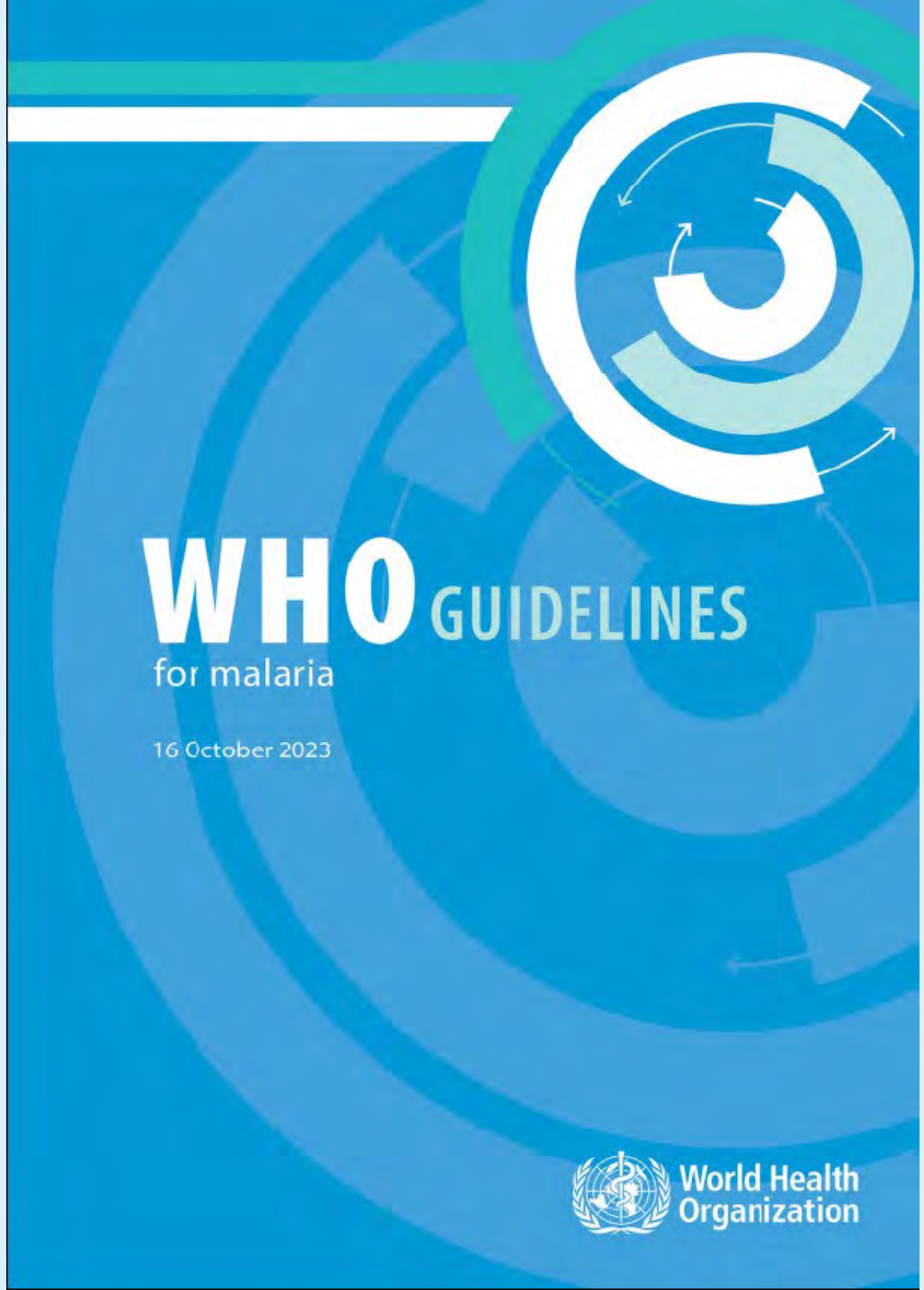
Nicotinic acetylcholine receptor competitive modulators

Neonicotinoids: clothianidin

GABA-gated chloride channel allosteric modulators

Meta-diamides: broflanilide

World Health Organization



Of note

Fig. 6. Non-inferiority margins for the primary end-point of mosquito mortality (pooled data for washed and unwashed nets) for the two studies assessing PermaNet® Dual as the candidate net compared to Interceptor® G2 as the active comparator

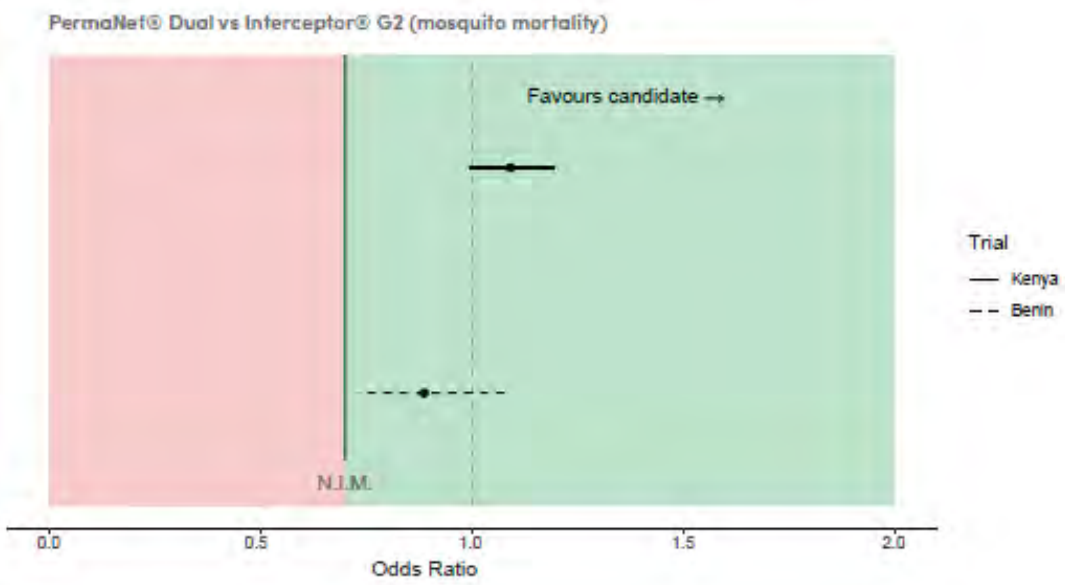


Table 15. Point estimates of pooled data from unwashed and washed nets for the respective products tested in the Benin trial

Outcome	Product	Role in study	Point estimate (%)	95% CI
Primary: Mortality (72 hours)	PermaNet® 2.0	Positive control	17.3	15.3–19.3
	Interceptor® G2	Active comparator	79.0	76.8–81.2
	PermaNet® Dual	Candidate	75.8	73.4–78.2
Secondary: Blood feeding	PermaNet® 2.0	Positive control	50.6	48.0–53.2
	Interceptor® G2	Active comparator	26.2	23.8–28.6
	PermaNet® Dual	Candidate	34.5	31.9–37.1

Table 18. Point estimates of pooled data from unwashed and washed nets for the respective products tested in the Kenya trial

Outcome	Product	Role in study	Point estimate (%)	95% CI
Primary: Mortality (72 hours)	PermaNet® 3.0	Positive control	56	54–57
	Interceptor® G2	Active comparator	65	64–67
	PermaNet® Dual	Candidate	68	66–69
Secondary: Blood feeding	PermaNet® 3.0	Positive control	6	6–7
	Interceptor® G2	Active comparator	10	9–11
	PermaNet® Dual	Candidate	12	11–13

"Although adopting the WHO-recommended fixed effects model resolved the situation in the present case, it was recognized that products that perform well (in terms of inducing high mortality) could end up being unable to demonstrate non-inferiority when a fixed OR is used, and that this challenge should be mitigated."
WHO Tech. Consultation Report, 2023

World Health Organization

Protocol Update

Incorporates recommendations from 2021 and 2023 technical consultations on comparative efficacy

Explicit recognition that "*at high mortalities, an OR of 0.7 imposes a near impossible condition for the candidate product to demonstrate non-inferiority (and requires very large sample sizes to obtain such narrow 95% CIs).*"

Recommend to WHO to modify methodology to:

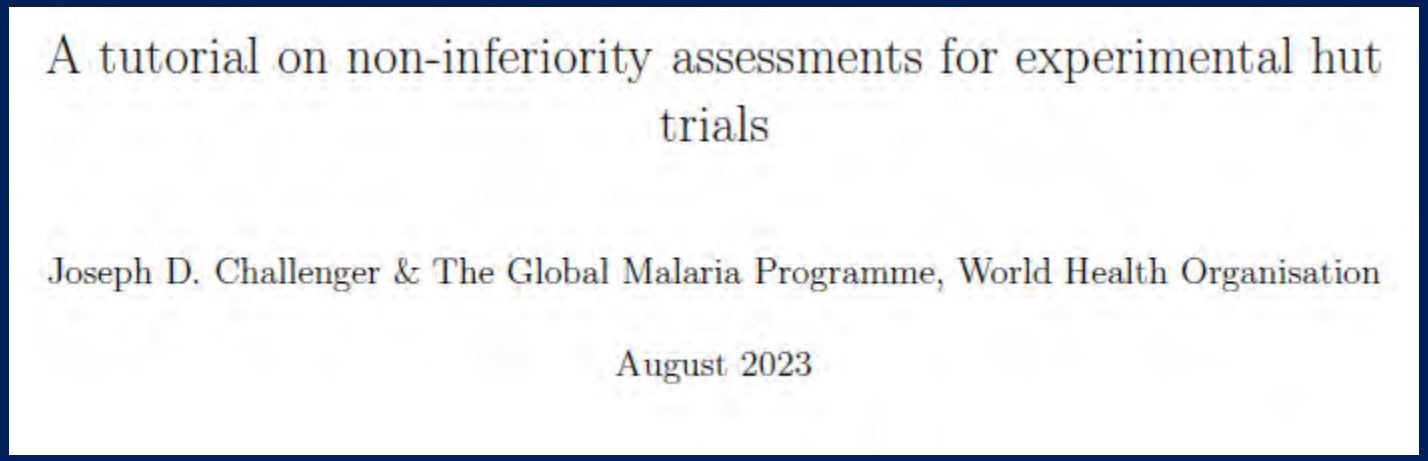
- preserve the use of non-inferiority as the sole decision-making approach
- preserve the use of an OR
- introduce an OR of the non-inferiority margin that varies depending on the percent mortality achieved by the first-in-class product

First-in-class product mortality (%)	Lower bound of candidate CI if non-inferiority mortality margin is 7%	Corresponding OR for a 7% non-inferiority margin
95	88	0.39
90	83	0.54
80	73	0.68
70	63	0.73
60	53	0.75
50	43	0.75
40	33	0.74
30	23	0.70

Modified approach uses a fixed percentage of 7% difference in mortality between the first-in-class product mortality and the lower bound of the candidate product's 95% CI to obtain the applicable OR

World Health Organization

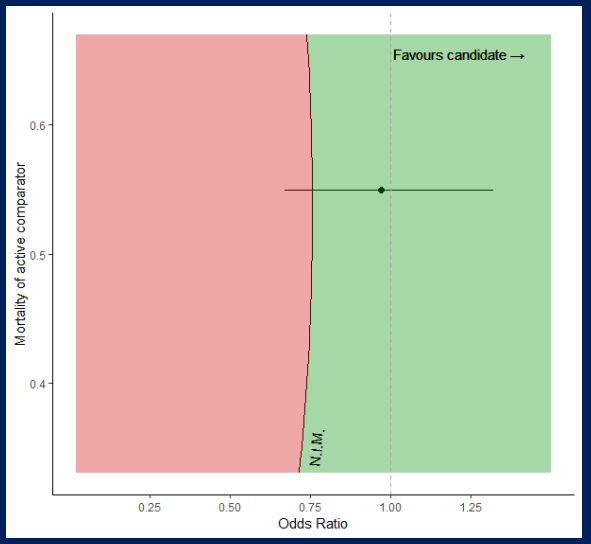
Tools developed



A tutorial on non-inferiority assessments for experimental hut trials

Joseph D. Challenger & The Global Malaria Programme, World Health Organisation

August 2023



Available from https://github.com/JDChallenger/WHO_NI_Tutorial

→ Video under development

Provides guidance on:

- Analysis in R or Stata
- Loading & summarizing datasets
- Making the non-inferiority assessment
- Summarizing and plotting the data
- Provides the analytical code to be used

World Health Organization

Evaluation & guidelines process update

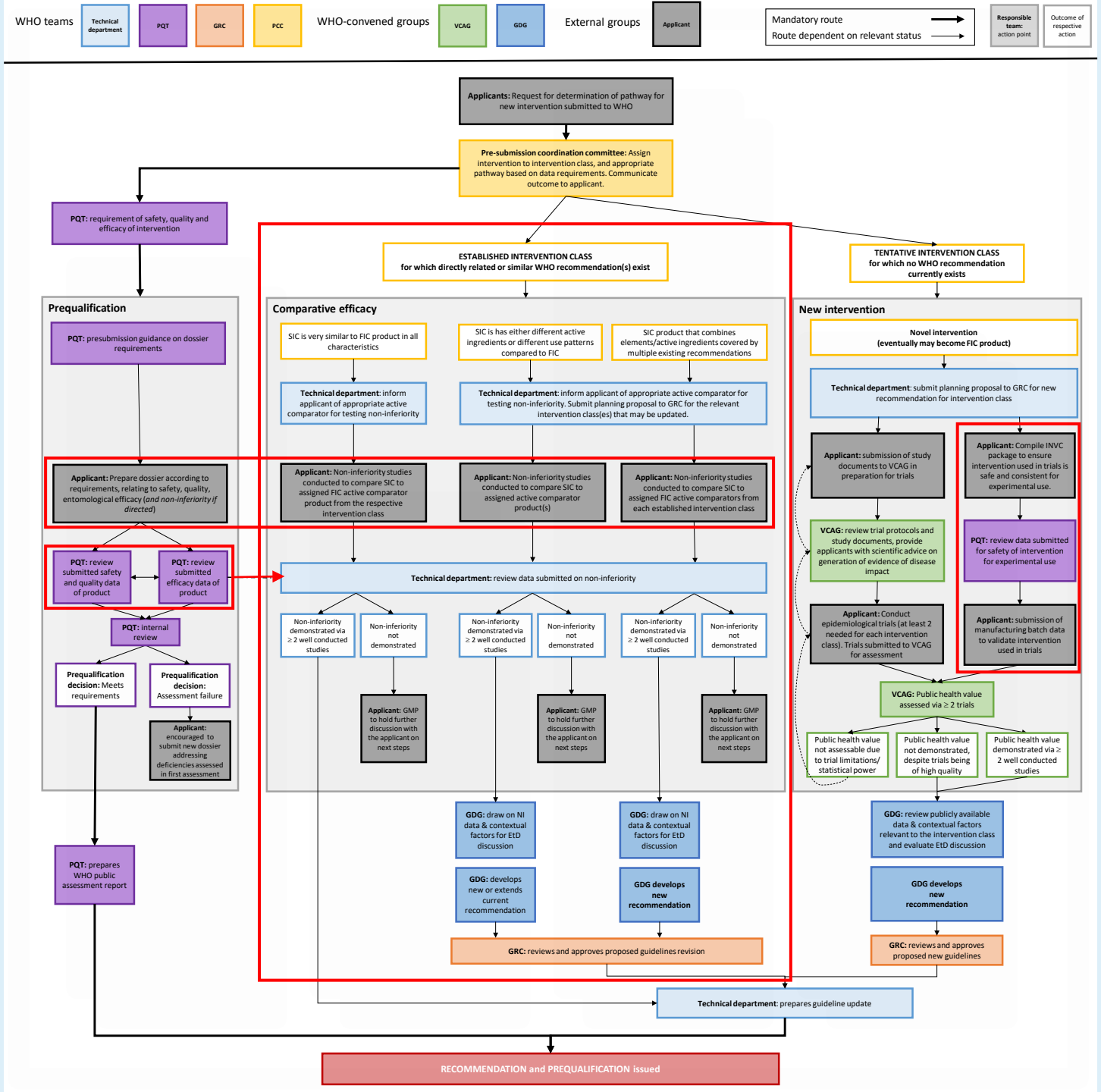
Discussions held with Guidelines Review Committee secretariat and chair to:

- Better define scope of recommendations
- Evidence-base and process to inform extension of existing recommendations or development of new ones

Discussions held with PQT-VCP on:

- Modifications to the evaluation process
- Roles and responsibilities

World Health Organization

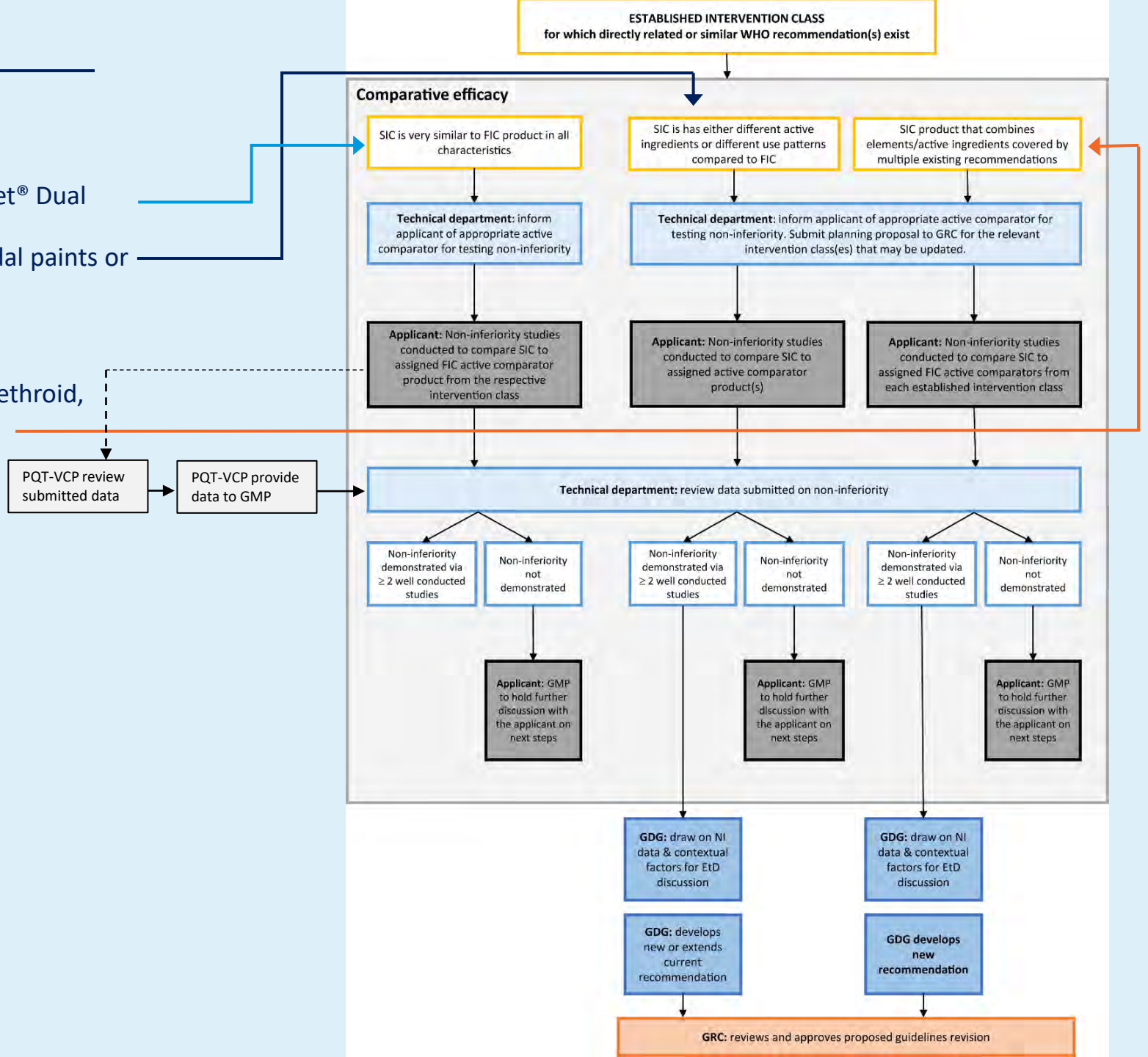


Examples

DuraNet Plus©, Yorkool® G3, PermaNet® Dual

SumiShield ®, Vectron T500, insecticidal paints or broflanilide treated ITN

Hypothetical example of ITN with pyrethroid, PBO, chlorfenapyr and pyriproxyfen



World Health Organization

Of note

Intervention Class ≠ WHO Recommendation

Update on work related to implementing a revised classification of ITN products

Background: May 2020, the classification of ITNs was revised into three classes summarized here:

2. ITNs designed to kill host-seeking insecticide-resistant mosquitoes for which a first-in-class product has demonstrated public health value compared to the epidemiological impact of pyrethroid-only nets.
3. ITNs designed to sterilize and/or reduce the fecundity of host-seeking insecticide-resistant mosquitoes for which a first-in-class product has demonstrated public health value compared to the epidemiological impact of pyrethroid-only nets.

MPAG April 2021 Meeting Report:
https://www.who.int/publications/i/item/9789240027350

≠

Conditional recommendation for , Moderate certainty evidence *

Pyrethroid-PBO ITNs (2022)

Pyrethroid-PBO ITNs instead of pyrethroid-only LLINs can be deployed for the prevention and control of malaria in children and adults in areas with ongoing malaria transmission where the principal malaria vector(s) exhibit pyrethroid resistance.

Strong recommendation for , Moderate certainty evidence *

Pyrethroid-chlorfenapyr ITNs vs pyrethroid-only LLINs (2023)

Pyrethroid-chlorfenapyr ITNs should be deployed instead of pyrethroid-only LLINs for prevention of malaria in adults and children in areas with pyrethroid resistance.

Conditional recommendation for , Moderate certainty evidence *

Pyrethroid-chlorfenapyr ITNs vs pyrethroid-PBO ITNs (2023)

Pyrethroid-chlorfenapyr ITNs can be deployed instead of pyrethroid-PBO ITNs for prevention of malaria in adults and children in areas with pyrethroid resistance.

World Health Organization

* https://www.who.int/teams/global-malaria-programme/guidelines-for-malaria

Of note

Intervention Class ≠ WHO Recommendation

Indicates the level at which epidemiological trials are (initially) needed

Are developed by considering a series of intervention-specific evidence-to-decision factors

Table 10.1. Factors that determine the direction and strength of a recommendation *

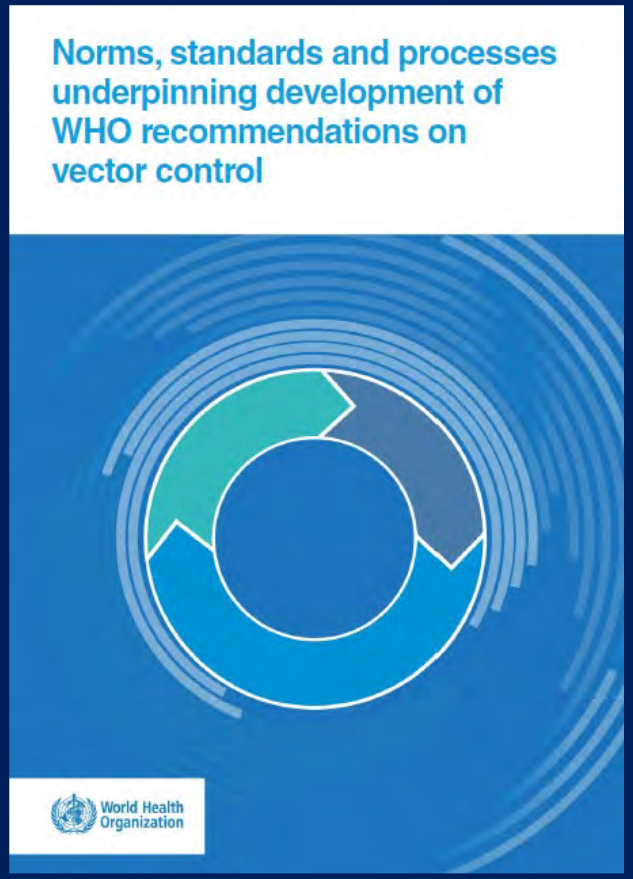
Factor	How the factor influences the direction and strength of a recommendation
Quality of the evidence	The quality of the evidence across outcomes critical to decision-making will inform the strength of the recommendation. The higher the quality of the evidence, the greater the likelihood of a strong recommendation.
Values and preferences	This describes how variable importance is assigned to health outcomes by those affected by them; how such importance varies within and across populations; and whether this importance is consistent or variable. The greater the uncertainty or variability there is about the values and preferences of people experiencing the critical or important outcomes, the greater the likelihood of a strong recommendation.
Balance of benefits and harms	This requires an evaluation of the absolute effects of both benefits and harms (or downsides) of the intervention and their importance. The greater the net benefit or net harm associated with an intervention or exposure, the greater the likelihood of a strong recommendation in favour or against the intervention.
Resource implications	This pertains to how resource-intense an intervention is, whether it is cost–effective and whether it offers any incremental benefit. The more advantageous or clearly disadvantageous the resource implications are, the greater the likelihood of a strong recommendation either for or against the intervention.
Priority of the problem	The problem's priority is determined by its importance and frequency (i.e. burden of disease, disease prevalence or baseline risk). The greater the importance of the problem, the greater the likelihood of a strong recommendation.
Equity and human rights	The greater the likelihood that the intervention will reduce inequities, improve equity or contribute to the realization of one or several human rights as defined under the international legal framework, the greater the likelihood of a strong recommendation.
Acceptability	The greater the acceptability of an option to all or most stakeholders, the greater the likelihood of a strong recommendation.
Feasibility	The greater the feasibility of an option from the standpoint of all or most stakeholders, the greater the likelihood of a strong recommendation. Feasibility overlaps with values and preferences, resource considerations, existing infrastructures, equity, cultural norms, legal frameworks, and many other considerations.

* WHO Handbook for Guidelines Development, 2nd Edition: https://www.who.int/publications/i/item/9789241548960

World Health Organization

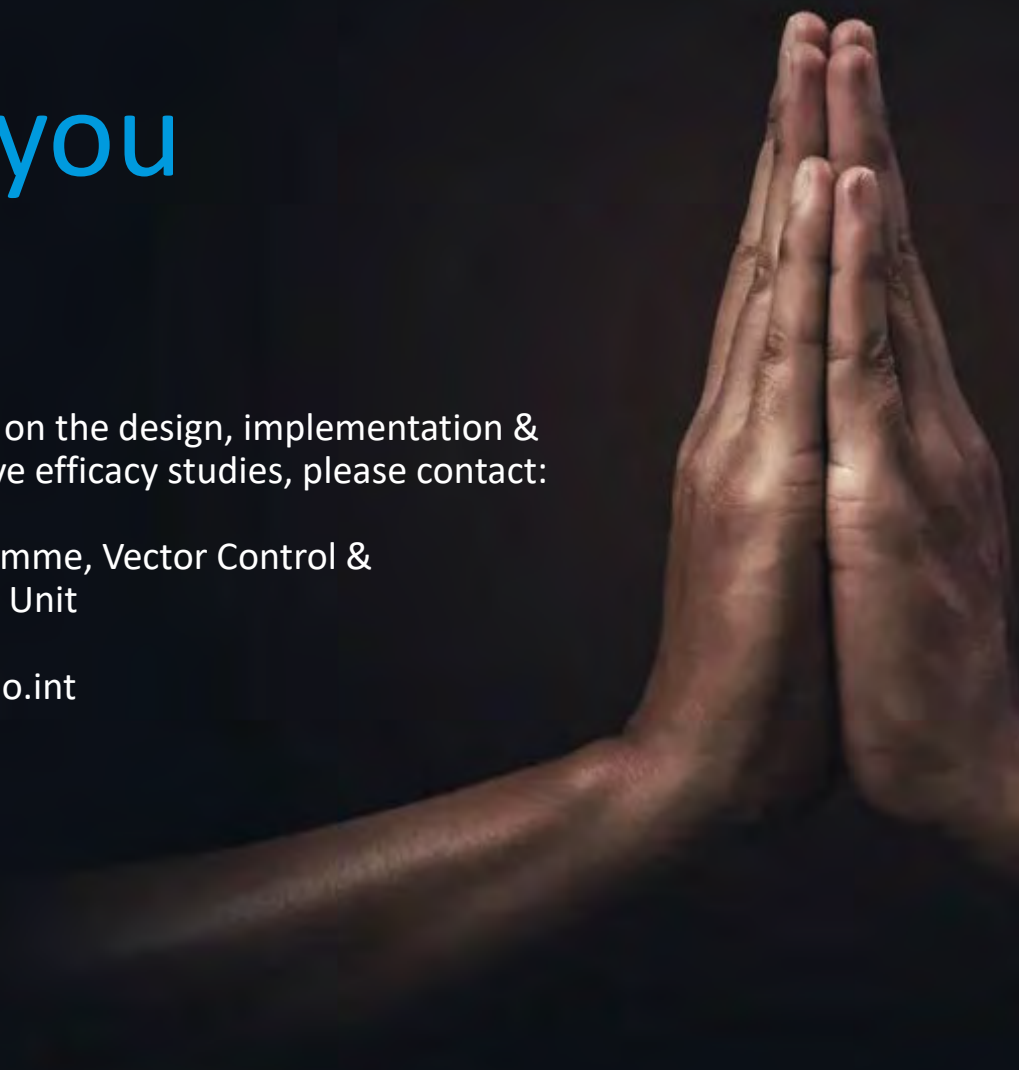
Next Steps

- Continue to participate in WHO's organization-wide alignment exercise of prequalification and normative processes.
- Update Norms, Standards and Process document including roles & responsibilities annex
- Investigate options to make comparative efficacy data publicly available (without having to publish meeting reports)
- Work with manufacturers and researchers to develop and evolve comparative efficacy testing methods for interventions other than IRS and ITNs. Note that this will be a phased approach intending to initially focus on spatial repellents.



World Health Organization

Thank you



For more information on the design, implementation & analysis of comparative efficacy studies, please contact:

Global Malaria Programme, Vector Control & Insecticide Resistance Unit

vc-noninferiority@who.int

Researchers

Manufacturers

Technical advisors

Modellers

Donors

WHO Colleagues

Malaria Policy Advisory Group Meeting
30 October–1 November 2023, Geneva, Switzerland
Background document for Session 5

World Health Organization

Evaluation of the "High burden to high impact" approach: lessons learned and future perspectives

Executive summary

Introduction

Malaria remains a public health problem in many countries, particularly in sub-Saharan Africa. According to the *World malaria report 2017 (1)*, 11 countries accounted for approximately 70% of the global estimated malaria case burden and 71% of global estimated malaria deaths. These countries are, in order of decreasing estimated number of cases: Nigeria (23%), Democratic Republic of the Congo (12%), Mozambique (5%), Uganda (5%) and Niger (4%), while Burkina Faso, Cameroon, India, Mali and the United Republic of Tanzania each accounted for 3% of the global burden of malaria cases. In 2021, with the ongoing increase in malaria cases, "High burden to high impact" (HBHI) countries accounted for 68% of all cases and 70% of deaths globally *(2)*.

While some countries have continued to reduce their malaria burden, reduction has stalled in many countries in the World Health Organization (WHO) African Region, which accounts for more than 90% of the burden of disease and for most of the increases in cases. These trends led to a call in 2018 to accelerate the reduction of malaria deaths and case incidence in the highest burden countries in Africa and in India.

The HBHI approach is a country-led response – catalysed by WHO and the RBM Partnership to End Malaria – to reignite the pace of progress in the global malaria fight. Four key mutually reinforcing pillars of the new HBHI response – political will to reduce malaria deaths, strategic information to drive impact, better guidance, policies and strategies, and a coordinated national malaria response – are underpinned by a recognition of the foundational supporting role played by the overall health system and the multisectoral response.

Objectives of the evaluation

The main objective was to evaluate and document the processes, lessons learned, best practices and challenges encountered in implementing the HBHI approach and to address gaps in the approach in Cameroon, Ghana, Mali and the Niger in order to better adapt and expand the approach to other countries. The outcomes of this evaluation in the four countries are complementary to the findings of same evaluation in the other 10 HBHI countries in Africa, which was completed in early 2022.

Specifically, the objectives were to:

- evaluate the country-level outcomes of applying the HBHI approach, identify best practices and barriers to success, and suggest course corrections for future actions;
- evaluate the global-level processes supporting the HBHI approach; and
- consolidate lessons learned and best practices, and set recommendations on the use of the lessons learned in the expansion of the HBHI approach to more malaria-endemic countries.

Methods

This evaluation focused largely on implementation of the HBHI approach at the country level. Mixed methods were used, consisting of an initial desk review, a survey, key informant interviews and four country case studies. Data were collected from late June to early September 2023. Overall, there were 60 key informants and 112 survey respondents from all four countries. In addition, eight global key informants were interviewed. Continuous data analysis was carried out throughout the evaluation process. Recorded in-depth interviews were transcribed verbatim and data matrixes constructed based on the main themes arising from the transcripts. Data were then organized into themes. All quantitative data were collected and entered into Google Forms. The data were then cleaned and analysed with Stata version 16. Simple frequencies and percentages were used for categorical variables. Basic cross-tabulation was done to establish relationships between variables and trends. Results were displayed in tables and graphs. Analysis consisted of consideration of each evaluation question, as well as findings and recommendations from both data sources to identify emergent findings.

Findings

Overall, 172 respondents were interviewed across the four countries, including 60 (34.9%) key informants and 112 (65.1%) survey respondents. In addition, a total of eight key informants were interviewed at the global level. The stakeholders that comprised the key informants varied by country, as this depended on who was involved in the HBHI approach in each country.

The roll-out of the HBHI approach was a collaborative effort, involving consultation, engagement and data-driven assessment. The implementation of the HBHI approach, however, has had varying degrees of success in improving programme performance at the country level.

Strategic information has been successfully integrated into programme reviews and data processes, providing a valuable framework for decision-making. The quality of data in-country has improved, which has enhanced the quality of national strategic plans. Overall, the approach has led to a broader recognition of the importance of data and subnational tailoring of interventions, reflecting positive outcomes. There has, however, been much less advancement with respect to the other HBHI components. The awareness and direct impact of the approach has yet to extend beyond the national level to the regional/provincial, district, subdistrict and community levels, with the coronavirus disease (COVID-19) pandemic further disrupting progress on implementation.

Global support from partners was immense in the preparatory activities for the initial stakeholder engagement, high-level political engagement and launch of the HBHI approach. Processes were in place to provide technical assistance for some of the pillars. WHO provided substantial technical assistance for pillar 2, in collaboration with funding and technical partner organizations. Following the launch of the HBHI approach in the countries, however, coordination meetings were not as effective as they should have been with regard to the form of the meetings and the participants involved, owing to emergence of the COVID-19 pandemic.

Table 1 presents the findings from each evaluation question, mapped against the evaluation objectives. Detailed findings are presented in the full report.

Table 1. Findings by objective and evaluation question

Evaluation question	Finding
Objective 1: Evaluate the implementation of the HBHI approach at the country level, including what has worked, what has not (the barriers to success) and what should be done differently. Each of the four pillars of the HBHI approach were assessed through this process (political will, strategic information to drive impact, better guidance, policies and strategies, and coordinated national malaria response). During this process, the lessons learned, best practices and challenges were documented.	
Evaluation question 1: To what extent have the objectives of the country programme been impacted by the HBHI approach?	• The implementation of the HBHI approach has had varying degrees of success in improving programme performance at the country level. • Strategic information has been successfully integrated into programme reviews and data processes, providing a valuable framework for decision-making. The quality of data has improved, which has enhanced the quality of national strategic plans. • Overall, the approach has led to a broader recognition of the importance of data and subnational tailoring of interventions, reflecting positive outcomes. • There has been much less advancement with respect to the other HBHI components. The awareness and direct impact of the approach has yet to extend beyond the national level to the regional/provincial, district, subdistrict and community levels. Most health managers below the national level have never heard of HBHI. • Implementation in all four countries was further disrupted by the COVID-19 pandemic, causing the initial momentum to slow considerably.
Objective 2: Evaluate the global-level processes supporting the HBHI approach.	
Evaluation question 2: To what extent have the global processes supporting the HBHI approach facilitated meeting the country's malaria objectives?	• Global support from partners was immense in preparatory activities for the initial stakeholder engagement, high-level political engagement and launch of the HBHI approach in all countries. • Processes were in place to provide technical assistance for some of the pillars. WHO provided substantial technical assistance for pillar 2, in collaboration with partner organizations. • Coordination meetings were not as effective as they should have been. In the view of participants, the meetings consisted more of presentations from countries than solid discussions. National malaria programmes were not involved in the meetings. In addition, the frequency of the meetings did not leave time for any concrete changes to be made. • While log frames were developed by in-country stakeholders to guide implementation, clear metrics to assess the effectiveness of the HBHI approach in yielding the desired outcomes were not established a priori. In addition, no monitoring and evaluation framework was embedded in the implementation.

Evaluation question	Finding
Objective 3: Consolidate lessons learned and best practices, and set recommendations on the use of the lessons learned in the expansion of the HBHI approach to more malaria-endemic countries.	
Evaluation question 3: How can examples of good practices and lessons learned from applying all components of the HBHI approach be adapted to different country contexts?	• Ongoing documentation of lessons should be facilitated. Lessons learned for subnational implementation should be drawn from all countries, particularly from unique countries with diverse subnational settings.
Evaluation question 4: How can examples of good practices and lessons learned from the HBHI approach inform the scale-up of this approach in additional malaria-endemic countries?	• A forum for peer learning and sharing of experiences should be created to enable countries to learn from each other. Such a collaborative platform would allow for the exchange of best practices and lessons learned, facilitating adaptation to local contexts. • Implementation guidelines that include best practices should be developed to guide implementing countries and any new malaria-endemic countries. • Some activities related to the HBHI pillars may require specific funding and additional human resources capacity in some cases. Country budgets should include those areas that have traditionally not been included.

Table 2. Lessons learned from implementation of the HBHI approach in the four countries

Components of the HBHI approach	What worked well	What did not work so well
General	• Perspective of national malaria programmes on malaria control/elimination broadened	• Perception of HBHI as a project and not an approach • Lack of ownership of the HBHI approach by some programme staff and national-level stakeholders • Expectation of additional funding that did not materialize • Lack of a forum for peer learning • Implementation of some aspects of HBHI not sustained • Momentum slowed with the emergence of the COVID-19 pandemic
Political will	• Concrete steps taken to ensure political will	• Translation of political will into domestic resources • End Malaria councils not yet set up in all countries
Strategic information for decision-making	• Marked improvement in the use of strategic information for decision-making, e.g. development of malaria strategic plans, subnational tailoring of intervention mixes based on the most current data • Improvement in the quality of data • Increased interest in data and their use • Integration of malaria data into national health management information systems in some countries where this was not the case before HBHI	• Repositories not yet functional in any of the four countries, though they have been initiated
Better guidance	• Development of new guidelines and update of existing guidelines	
Programme coordination	• Increased country stakeholder involvement	• Limited or no dissemination of the HBHI approach to subnational levels
Multisectoral action		• Position of the national malaria programme within the health sector in general makes it challenging for it to get other sectors to contribute effectively • Political will has yet to facilitate action by all sectors • Lack of understanding of the malaria problem by other sectors • Inadequate resources and unclear ownership and stewardship

Components of the HBHI approach	What worked well	What did not work so well
Integrated health system	• Some existing areas of integration in the health system, such as with the reproductive, maternal, newborn and child health programmes for delivering some interventions that continue to work well	• Very few innovative areas for further integration

Conclusions

The implementation of the HBHI approach has had varying degrees of success in improving programme performance at the country level. Notably, there have been significant advancements in certain areas, particularly in strategic information and political will. In these areas, time and effort have yielded obvious improvements. The quality of data has improved, and the quality of national strategic plans has also been enhanced. The approach's overall impact has been a broader recognition of the importance of data and subnational tailoring of interventions, reflecting positive outcomes.

The extent of progress, however, has been slightly constrained by challenges related to coordination, the development of better guidance and policies locally, and the ability to effectively engage other sectors and critical non-health stakeholders in the fight against malaria. These areas have shown comparatively less progress. Translating the expression of political will into an increase in domestic financing has also proven to be a bottleneck. Roll-out of the HBHI approach has still not moved beyond the national level, with a general lack of awareness observed below the national level.

In sum, while the HBHI approach has played a significant role in improving malaria programme performance in countries, measuring its impact can be challenging without defined performance metrics. To determine the success of the HBHI approach, a monitoring and evaluation framework should have been included and clear metrics established a priori to assess the effectiveness of the approach in yielding the desired outcomes.

Recommendations

The following recommendations were synthesized from recommendations provided by study participants from the four countries and by global respondents:

1. **Ensure a shift in mindset from viewing HBHI as a project to seeing it as an approach:** Fundamentally, there is a need for a shift in perspective, moving away from viewing HBHI as a project towards seeing it as an approach.
2. **Rethink how to translate expressed political will into increased domestic resources**: Most countries have not been able to leverage expressed political will to achieve a tangible increase in domestic resources for malaria elimination efforts. There is a need to rethink the implementation of this pillar and identify solutions to alleviate this bottleneck.
3. **Tailor HBHI efforts to each country's specific context and strengths:** These observations underscore the importance of making every effort to tailor the HBHI approach to each country's specific context and strengths. While some elements may be easier to address in certain countries, the overall success of the approach hinges on it being a comprehensive approach that aligns with the unique needs and capacities of each participating country.
4. **Build the capacity of national malaria programme managers and health managers to lead the effort:** The capacity to lead and implement the different components of the HBHI approach

requires a paradigm shift in mindset with respect to how malaria programmes are traditionally run and a move out of comfort zones into areas that are not the norm for programme staff, their stakeholders and political authorities. This, therefore, requires some level of orientation and capacity-building, particularly in some of the components that may be new, such as political will, multisectoral action, and, in some cases, coordination of all actors and partners and their contribution to achievement of the set objectives. Of particular importance is the expansion of the skill set to include softer skills, such as social and behavioural sciences and political analytics.

5. **Additional human resources allocation and funding are vital:** Human resources allocation and some specific funding are vital for efficient execution and coordination of the approach. Some of the areas of HBHI activities have not traditionally been included in country budgets and this needs to be reviewed. The specific funding could be included in existing funding streams or supported by partners once identified. Additional human resources may be needed in the short to medium term in some cases.

6. **Decentralize the HBHI approach and ensure that it is deployed at all other subnational levels:** While high-level political involvement is deemed essential at the national level, it is important to find ways to decentralize the HBHI approach down to all levels in order to maximize the impact of the approach. Most subnational health managers had never heard of HBHI. In essence, every effort should be made to deploy the HBHI approach down to the regional or provincial, district and community levels, as each of these subnational levels is actually a microcosm of the next higher level. This shift has to be intentional and requires funding to accomplish. This will then mainstream the thinking and approach in the efforts of all the different levels to control/eliminate malaria.

7. **Set up a forum for sharing experiences and lessons learned**: Such a forum was universally thought to be a critical component that was absent from the roll-out and implementation, which created a gap among implementers in terms of where to seek advice and support as implementation continued and bottlenecks were identified. It was thought that it would be better if the forum could be country-led but facilitated and supported by one of the global partners.

8. **Develop implementation guidelines to provide guidance to countries:** Further to the call to ensure orientation of all programme implementers and the set-up of a forum to share experiences, the importance of developing guidelines to provide guidance to countries based on the experiences gathered from the roll-out in the initial set of countries was emphasized.

9. **Build a monitoring and evaluation framework into the HBHI approach:** A monitoring and evaluation framework needs to be built into the implementation of the HBHI approach up front, so that these activities are done at regular intervals, with the results feeding into an improvement in the implementation and course correction where necessary.

10. **Effective multisectoral action will ride on the back of political will at the highest level and should be a key outcome:** Multisectoral action will ride on the back of political will at the highest level and should be one of the expected outcomes of the political will pillar. The position of the national malaria programme within the health sector structure makes it challenging to get other sectors outside of health to contribute effectively at the highest level. Involvement of representatives in technical working groups of the national malaria programmes may ease the path through which effective engagement may be achieved; however, political will at the highest level in the country is needed to achieve maximum impact.

In sum, the lessons learned from the HBHI approach underscore the importance of integrating strategic information into regular processes, defining clear performance metrics and shifting from a project-centric view to an approach that adapts interventions to specific contexts. In addition, there is recognition of the need to seize opportunities for integrated health management during crises such as the COVID-19 pandemic.

References

1. World malaria report 2017. Geneva: World Health Organization; 2017 (https://iris.who.int/handle/10665/259492, accessed 25 October 2023).
2. World Health Organization, RBM Partnership to End Malaria. HBHI response element: country engagement. Presentation at the iCCM technical consultation, Addis Ababa, 25 July 2019 (https://www.childhealthtaskforce.org/sites/default/files/2020-04/HBHI%20Response%20Element_Country%20Engagement_iCCM%20Technical%20Consultation_07.2019.pdf, accessed 25 October 2023).

Update on the High burden to high impact (HBHI) approach

MPAG meeting 30 Oct-01 Nov, 2023

Dr Maru Aregawi

Unit Head, High Burden High Impact (HBI)
Global Malaria Programme (GMP),

World Health Organization

RBM Partnership To End Malaria

Introduction: Progress towards GTS targets

Endorsed by 74th WHA in 2021

Global **Malaria** Programme

WHA74.9

1. RECOMMITS to the goal of malaria eradication and affirms that this goal will be incorporated into the post-2030 iteration of the global technical strategy for malaria

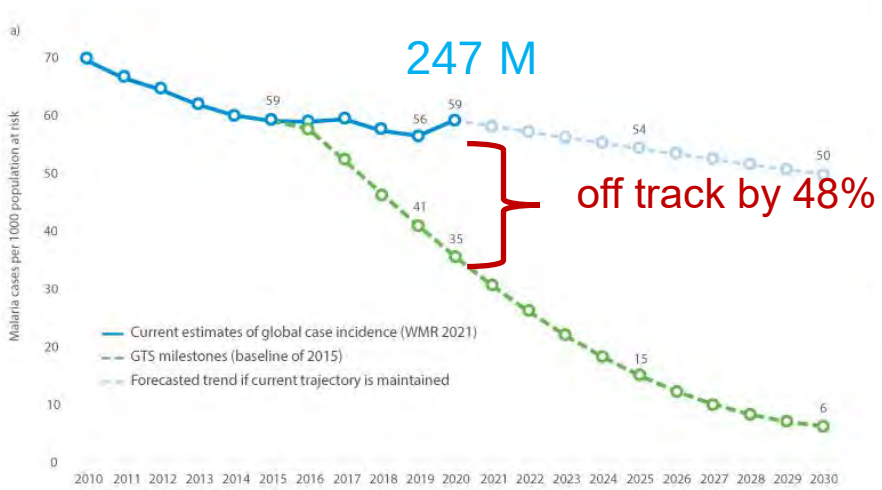
World Health Organization

Number of endemic countries 93➔ 84

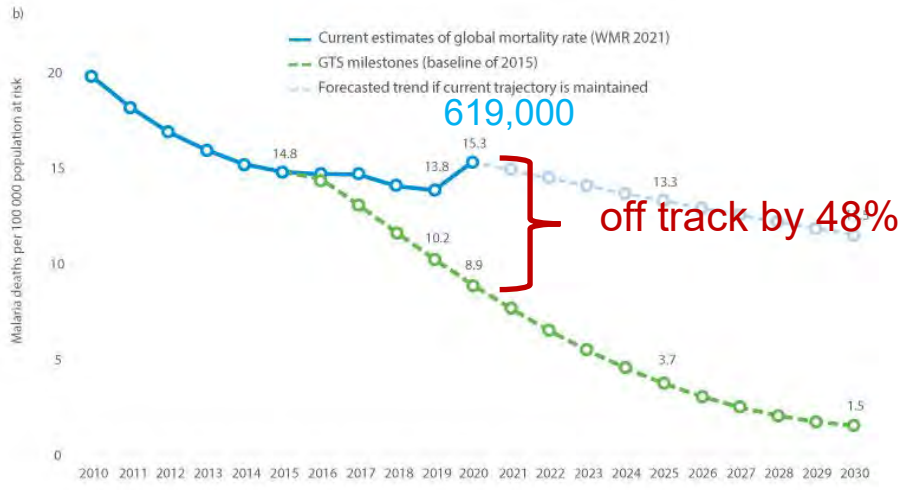
(WMR, 2022)

Progress towards GTS targets is substantially off track (WMR, 2022)

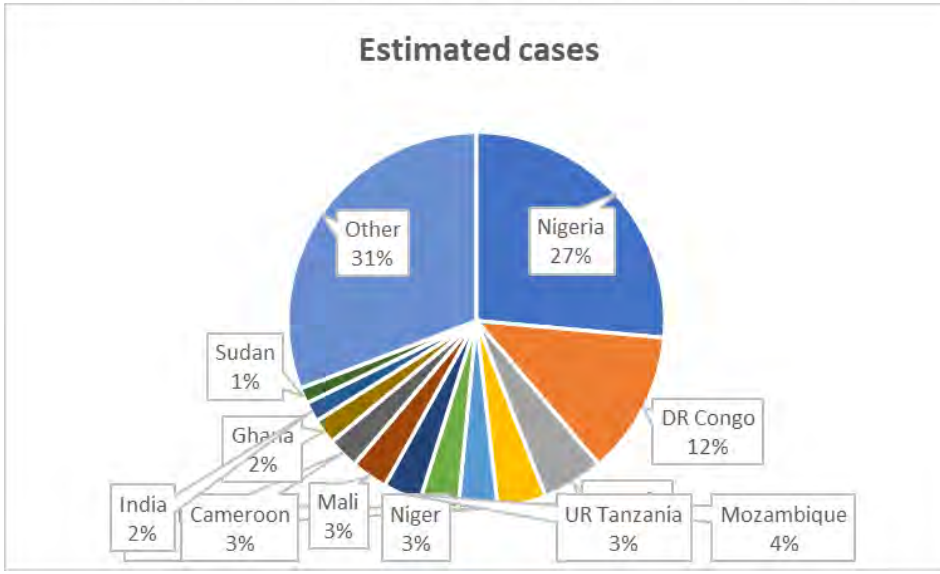
A) Cases



B) Deaths



- Countries of the WHO African Region accounted for >95% of cases and deaths
- 11 HBHI countries: 71% of global mortality
- Five countries: 52% (Nigeria, DRC, Uganda Mozambique, Burkina Faso)



HBHI: The goal, objective and approaches

The problem: Malaria, in the high burden countries of Africa, remains as the major cause of death but continues to be the "**the abnormal" accepted as "normal**". Global progress stalled due to slow progress in the 10 Countries of Africa accounting to 70% of the global burden. India has progressed since 2018.

Goal: To get the world back on track to achieve GTS milestones by 2025 and sustain gains to reach the GTS goals by 2030.

Core objective: accelerate reduction of malaria mortality while reducing malaria incidence through concerted political will, use of data for action, better guidance and coordination, founded by strong health system and multisectoral approach.

Principles

- Country-owned, Country-led
- Better coordination
- Partners commitment
- Increased domestic financing

Countries with the highest burden of malaria (80%), WMR 2021

S.N	Country	Estimated cases	Estimated deaths	%	Incidence / 1000	HBHI
1	Burkina Faso	8150690	19979	3%	401.0893	1st
2	Benin	4707522	10123	2%	398.9036	2nd
3	Mali	7238665	19316	3%	368.2296	1st
4	Liberia	1810880	4601	1%	366.7699	2nd
5	Central Africa	1622774	5079	1%	341.9837	2nd
6	Niger	7845520	17435	3%	336.5628	1st
7	Sierra Leone	2617968	8054	1%	335.0696	2nd
8	Democratic R	29036471	82511	12%	334.5579	1st
9	Mozambique	10007802	23766	4%	329.5721	1st
10	Guinea	4196430	10215	2%	328.5842	2nd
11	Nigeria	64677959	199689	27%	321.8392	1st
12	Burundi	3506219	5822	1%	304.0801	2nd
13	Cote d'Ivoire	7571801	15913	3%	294.433	2nd
14	Uganda	12982098	21699	5%	293.251	1st
15	South Sudan	3211331	7431	1%	290.2999	2nd
16	Cameroon	6900814	14841	3%	266.6838	1st
17	Angola	8268572	15989	3%	259.8113	2nd
18	Equatorial Gu	337892	674	0%	249.1862	2nd
19	Rwanda	2986047	3046	1%	236.4823	2nd
20	Malawi	4370301	7165	2%	234.5998	2nd

High Burden Hig

ership To End Malaria

World Health Organization

Update on HBI activities

2023 progress to date

- HBHI country updates
- Support countries in NSPs, MPRs MTR, GF proposals
- Support countries in responding to epidemics and emergencies
- HBHI Evaluation
- Draft malaria control in Emergencies manual
- Implementation of 1,7 mRCTR operational research (funded by UNPDF)

Priorities for next quarter

- Stakeholders' review meeting for finalization of malaria control in emergencies (5-8 Dec, 2023)
- Finalization of 1,7 mRCTR operational research (funded by UNPDF) in 3 countries (Senegal, Zambia and Tanzania), Demanding No-Cost Extension for Burkina Faso (due to conflict)
- Revision of the epidemic preparedness and response (within the surveillance manual)

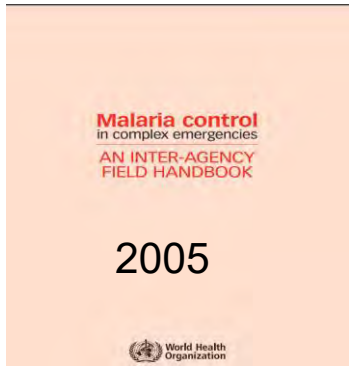
Priorities for 2024

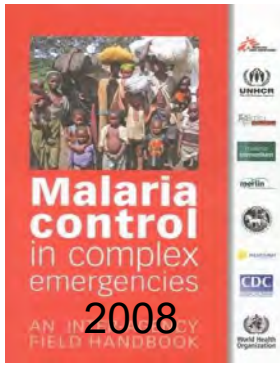
- Printing of malaria in emergency manual
- HBHI country updates
- Country support and strengthening
- SOP for mortality mapping and accelerated response at district level

Update: Malaria control in emergencies manual

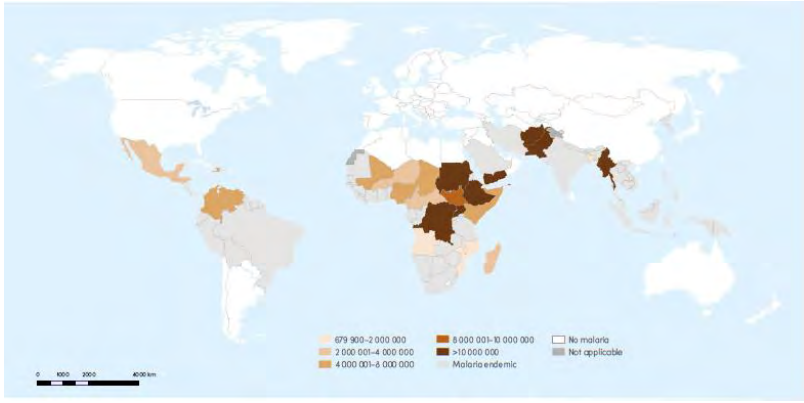
Malaria control in emergencies
Third Edition (draft) will address

- Globally, 340 million people affected by disasters and humanitarian crises requiring assistance in 2023, projected to increase.
- 122 M people in 21 endemic countries
  - Afghanistan, Colombia, the Democratic Republic of the Congo, Ethiopia, Mali, Myanmar, Nigeria, Somalia, South Sudan,
- the Sudan, Uganda and Yemen.
- 9 of the 12 HBHI countries experiencing some form of conflict and emergencies
- Update policies and strategies
- Best practices and lessons learnt
- Integrate malaria with the WHO Health Emergencies
  - Incident Management System





People in humanitarian need in malaria endemic countries (WMR 2022)



Update: 1,7 mRCTR operational research in moderate-high transmission

UR Tanzania: Phase 1

- 1, 7 malaria reactive community-based test and respo
  (1,7mRCTR)
- **Objective: Assess impact of** 1,7mRCTR

- **Findings:**
  - Malaria parasite prevalence by 81% in intervention areas (26%4.9%) or
  - 66% additional reduction of malaria parasite rate compared to control areas

Malaria Journal
BMC
Malar J. 2020; 19: 292.
Published online 2020 Aug 14. doi: 10.1186/s12936-020-03363-w
PMCID: PMC7429894
PMID: 32799857

Effectiveness of the innovative 1,7-malaria reactive community-based testing and response (1, 7-mRCTR) approach on malaria burden reduction in Southeastern Tanzania

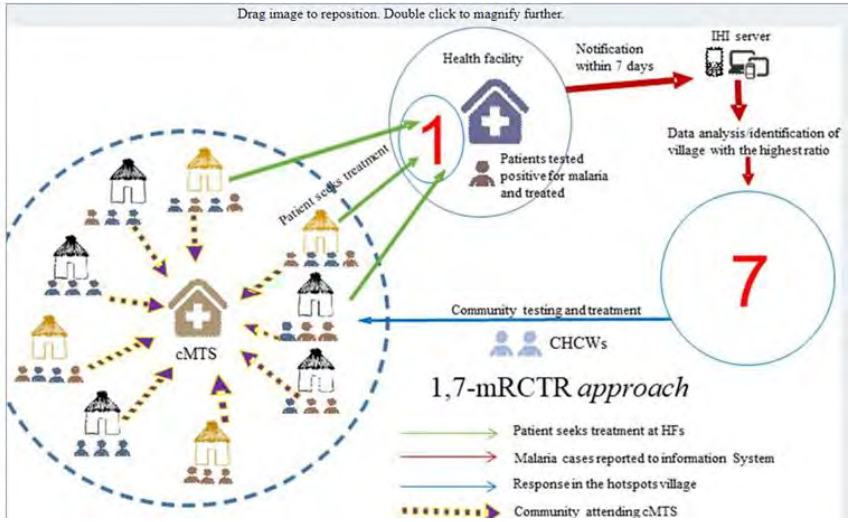
Yeromin P. Mlacha,[2,3,4] Duoquan Wang,[1] Prosper P. Chaki,[2] Tegemeo Gavana,[2] Zhengbin Zhou,[1] Mihayo G. Michael,[2] Rashid Khatib,[2] Godlove Chila,[2] Hajirani M. Msuya,[2] Exavery Chaki,[2] Christina Makungu,[2] Kangming Lin,[5] Ernest Tambo,[6] Susan F. Rumisha,[8] Sigsbert Mkude,[2] Muhidin K. Mahende,[2] Frank Chacky,[7] Penelope Vounatsou,[3,4] Marcel Tanner,[3,4] Honorati Masanja,[2] Maru Aregawi,[9] Ellen Hertzmark,[10] Ning Xiao,[1] Salim Abdulla,[2] and Xiao-Nong Zhou[1]

UR Tanzania: Phase 2

- 2019-2022
- Findings: 37% additional reduction in intervention areas

Ongoing project: UNPDF (~1.7 million USD)

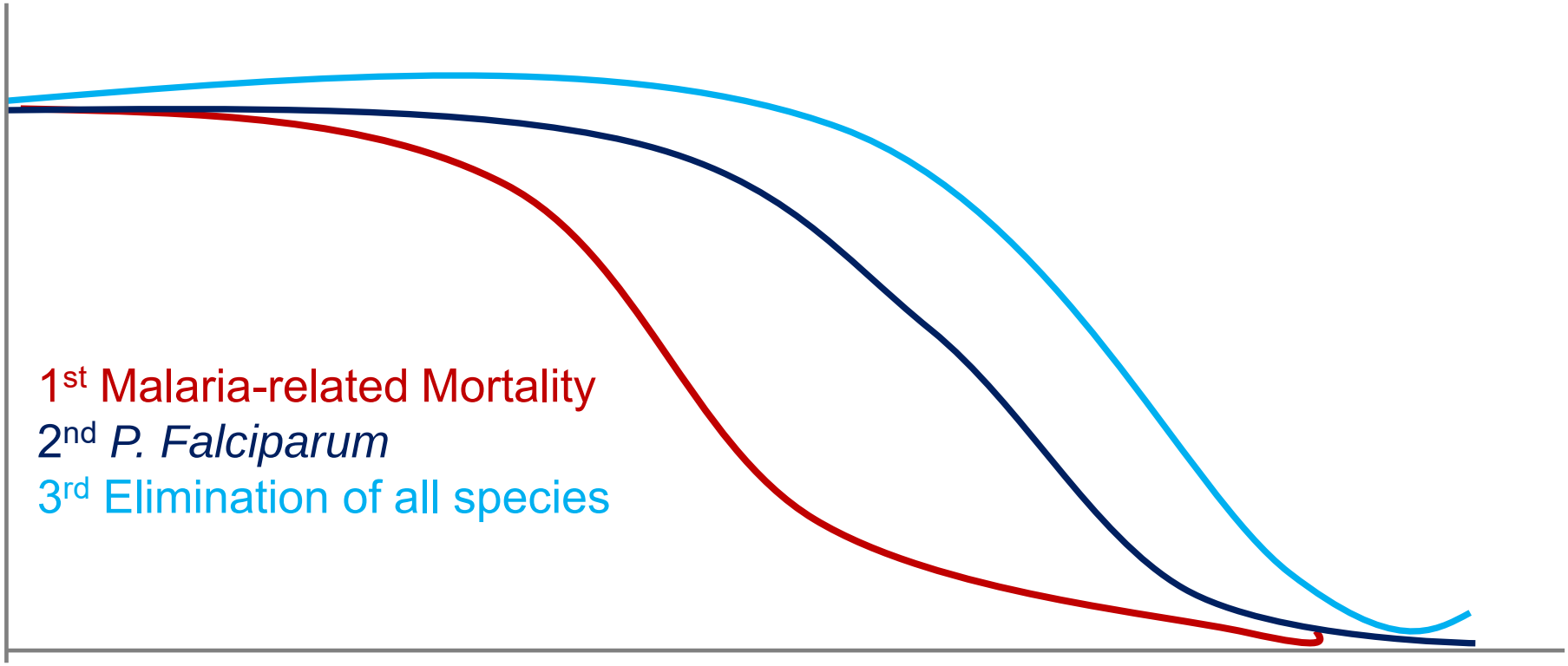
- Objective: 1,7mRCTR and Multisectoral approach
- 2021-2023
- Burkina Faso, Senegal and Zambia (1,7mRCTR)
- UR Tanzania to monitor post-project resurgence



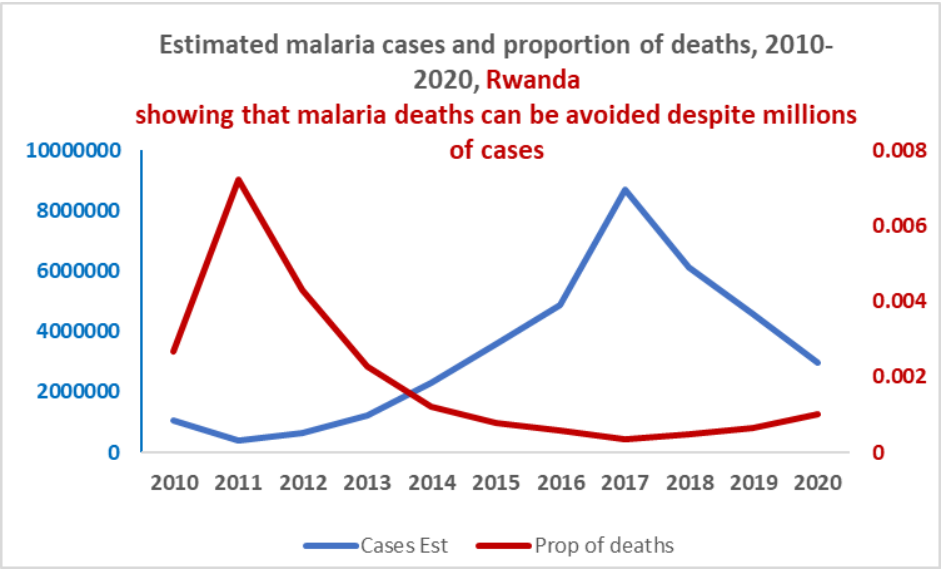
Strategic shift for HBHI

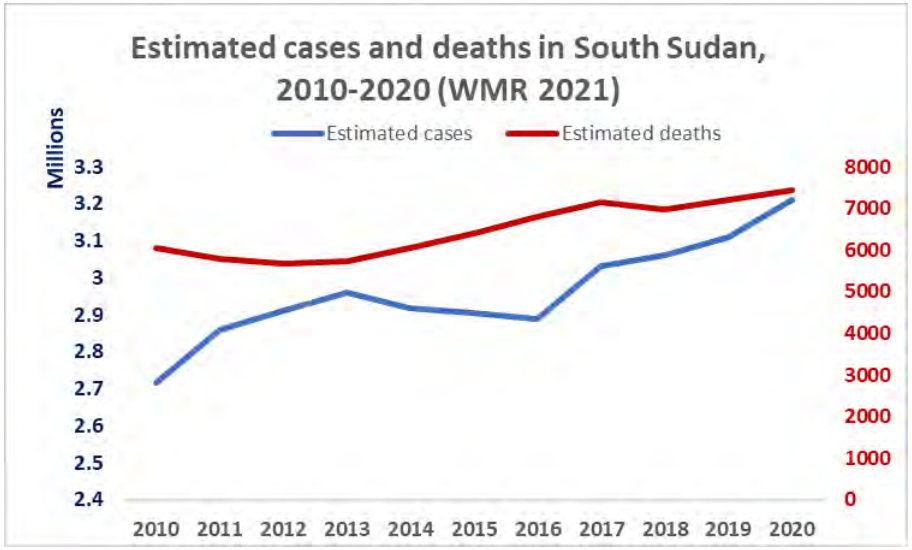
- Focus on accelerated malaria mortality reduction
- Change the mindset of equating the targets for malaria mortality and incidence in HBHI countries
  - **Infection ≠ Incidence ≠ severe malaria ≠ malaria death**
- With existing tools & systems:
  - It is possible to end malaria-related death
    - Refocusing, policy, leadership and commitments, health system, surveillance & local data use, CHWs.
  - Nearly impossible to interrupt transmission and avoid malaria episodes because of the:
    - epidemiology in SSA
    - sub-optimal effect of the preventive tools

Impactful public health strategy for malaria



The case of Rwanda and South Sudan Malaria





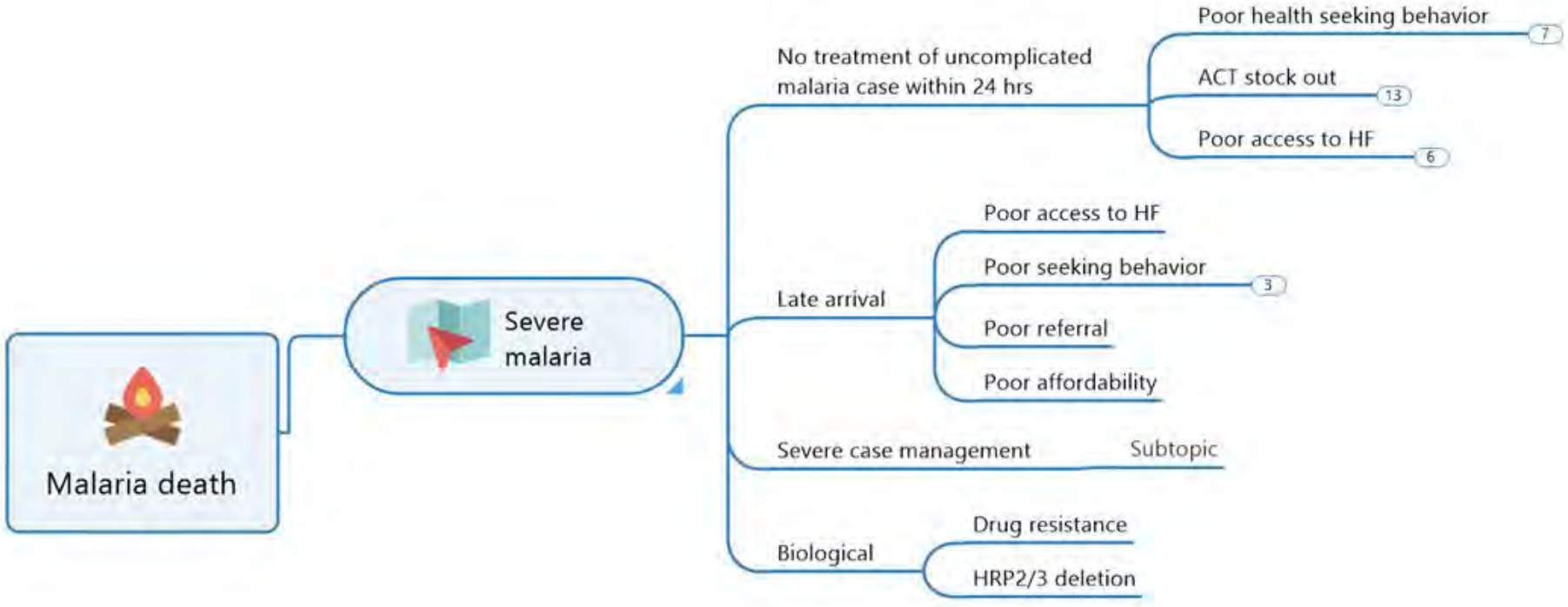
1. Strong political commitment, leadership and accountability
2. Strong health system: Early diagnosis and treatment through strong PHC and CHW
3. Strong surveillance and referral systems
4. Functional community-Based Health Insurance (CBHI)

1. Strengthen political commitment and accountability
2. Strengthen the health system: Early diagnosis and treatment through strong PHC and CHW, Chronic complex emergency
3. Strengthen community networks, Integrated approach
4. Strengthen surveillance and referral systems

SOP for critical mapping for accelerated malaria mortality reduction

Objective: **To enable NMCPs in moderate-high transmission settings identify and understand the various factors contributing to malaria-related deaths.**

By recognizing these social, technical, and operational factors, the aim is to prioritize and take action to mitigate the primary drivers of malaria mortality within their healthcare systems.



HBHI Evaluation objectives

EQ1	Global Implementation Processes: To what extent has the process of global HBHI implementation facilitated improved malaria programme engagement with partners?
EQ2	Impact on Country Level Performance: To what extent has HBHI implementation led to improved performance at the country level?
EQ3	Scaling up all 4 Elements: How can examples of good practices and lessons learned from HBHI implementation inform the scale up of all four elements?
EQ4	Scaling up HBHI to additional countries: How can examples of good practices and lessons learned from HBHI implementation inform the scale up to additional countries?

HBHI Evaluation

Two batches

- RBM sponsored (2022): Six countries (Burkina Faso, DR Congo, Mozambique, Nigeria, Tanzania, Uganda)
  - Presented in previous MPAG meeting

- WHO sponsored (2023): 4 countries (Cameroon, Ghana, Mali, Niger)
  - July-October 2023
  - Led by Prof Evelyn Ansah (Overall coordinator)
    - Dr Philippe Nwane: Cameroon
    - Dr Mohamed Traore: Mali
    - Dr Valentine BATAMU KAMANDA: Ghana
    - Dr Goubekoy Bawan Allah: Niger

Evaluation of the HBHI Approach:

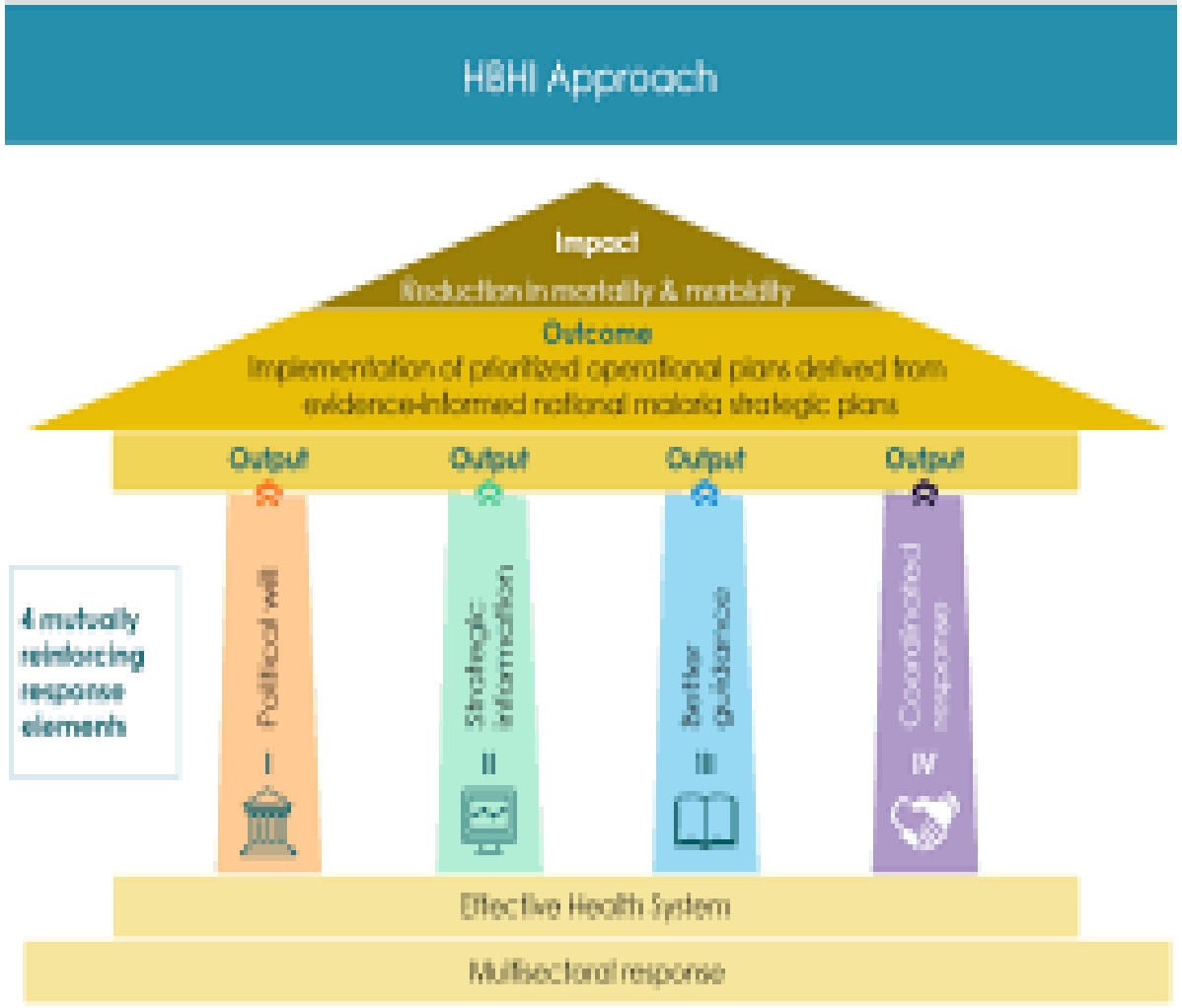
Lessons Learnt and Future Perspectives

INTRODUCTION

Background

- Malaria remains a public health problem in many countries, particularly in sub-Saharan Africa
- Eleven (11) countries accounted for approximately 70% of the global estimated malaria case burden and 71% of global estimated malaria deaths
- In 2021, HBHI countries accounted for 68% of all cases and 70% of deaths globally
- The HBHI approach is a country-led response – catalysed by WHO and the RBM Partnership – to reignite the pace of progress in the global malaria fight.

Pillars of the HBHI Approach



I **Political will** to reduce malaria deaths

II **Strategic information** to drive impact

III **Better guidance,** policies and strategies

IV A **coordinated** national malaria **response**

Underpinned by Multisectoral Action and a Functioning Health System

World Health Organization

RBM Partnership To End Malaria

Purpose and Scope of Evaluation

Purpose of the Evaluation

- Document the lessons learned, best practices and challenges encountered in implementing the high burden to high impact (HBHI) approach
- Consider how the approach can be further adapted based on the lessons learned and best practices observed address the current context and challenges and expand to other countries
- Improve the global implementation of the approach

Scope of the Evaluation

- This is not an evaluation of country performance, per se
- It focuses on the process of the HBHI implementation as an approach
- It assesses what has worked well and what did not work as planned, documents lessons learnt, success stories and challenges to improve the approach further in the implementing countries.

Objectives of the Evaluation

Main objective

- Evaluate and document the processes, lessons learned, best practices and challenges encountered in implementing the HBHI approach and to address gaps in the approach in Cameroon, Ghana, Mali and the Niger in order to better adapt and expand the approach to other countries.

The outcomes of this evaluation in the four countries are complementary to the findings of same evaluation in the other 10 HBHI countries in Africa, which was completed in early 2022.

Specific objectives

- Assess and document country-level outcomes of applying the HBHI approach, identify best practices and barriers to success, and suggest course corrections for future actions;
- Assess the global-level processes supporting the HBHI approach; and
- Consolidate lessons learned and best practices, and provide recommendations for the use of the lessons learned in the expansion of the HBHI approach to more malaria-endemic countries.

Evaluation Questions

- **EQ1: IMPACT ON COUNTRY LEVEL PERFORMANCE**: To what extent have the objectives of the country programme been impacted by the HBHI approach?

- **EQ2: GLOBAL IMPLEMENTATION PROCESSES**: To what extent have the global processes supporting the HBHI approach facilitated meeting the country's malaria objectives?

- **EQ3: APPLYING ALL COMPONENTS OF THE APPROACH**: How can examples of good practices and lessons learned from applying all components of the HBHI approach be adapted into different country contexts?

- **EQ4: SCALING UP HBHI TO ADDITIONAL COUNTRIES**: How can examples of good practices and lessons learned from the HBHI approach inform the scale-up of this approach to additional malaria-endemic countries?

Methods

Desk Review

Key Informant Interviews (KII) - Global

Country Case Studies incl KIIs and Electronic Survey – Ghana, Cameroun, Mali, Niger

Methods

- Initial Stakeholder Meetings to brief them about the evaluation
- In-depth interviews were recorded and transcribed verbatim.
- Qualitative data were organized into themes.
- Quantitative data were entered into Google Forms, cleaned and analyzed with STATA version 16
- Informed Consent was sought from all participants
- Use of standard tools by country consultants for country case studies
- Regular check-ins carried out between the global and in-country consultants to discuss and agree on how to approach any areas of challenge

Stakeholder Meeting

World Health Organization

RBM Partnership To End Malaria

Study Participants

NATIONAL

- Malaria Programme Manager
- NMCP Staff
- Other National Health Staff
- NGO/CSO Reps
- Devlopment Partner
- Implementing Partner
- Traditional Leader
- Parliamentarian/Mayor
- Researcher
- Academician
- Local Government Staff

SUB NATIONAL

- District Director of Health Services
- District Malaria Focal Point
- District Monitoring & Evaluation Manager
- District Surveillance Officer
- District Health Information Officer

KEY FINDINGS AND LESSONS LEARNT

Participants

- Desk review involved over 81 documents
- Overall, 172 respondents interviewed across the 4 countries
  - 60 (34.9%) key informants
  - 112 (65.1%) survey respondents
- Additionally, a total of eight (8) key informants were interviewed at global level
- Overall, in-country respondents comprised 18.6% (32) females**; All survey respondents in Mali male;** 25.0%(2) of 8 global respondents were female

Country	Key Informant Interview	Survey
Ghana	23	30
Cameroun	13	32
Mali	7	37
Niger	17	13
Total	**60 (34.9%)**	**112 (65.1%)**

Country HBHI Initiation and Roll-out

- All 4 countries launched the HBHI Approach as follows:

  Cameroun - May 2019; Ghana - June 2019; Niger - September 2019. **Mali - April 2021**

- **This means 10 to 12 months implementation period before the start of COVID-19 Pandemic**
- The engagement process facilitated by the Global Malaria Programme (GMP) of the WHO, and RPM Partnership to End Malaria enabled countries carry out a holistic self-assessment and stakeholder engagement at national level
- High-level officials participated in official launches and national commitments.
- Stakeholder engagement on malaria control involved various stakeholders in each cou

"Before this meeting, it didn't really seem like, people accepted that they are a part of the whole fight against malaria or when it comes to malaria, they really had critical role to play but just that meeting really gave that first recognition that we are all partners in it and if we really need to move ahead then all the contributions from these various partners are really needed" **(Staff of NMEP, Ghana)**

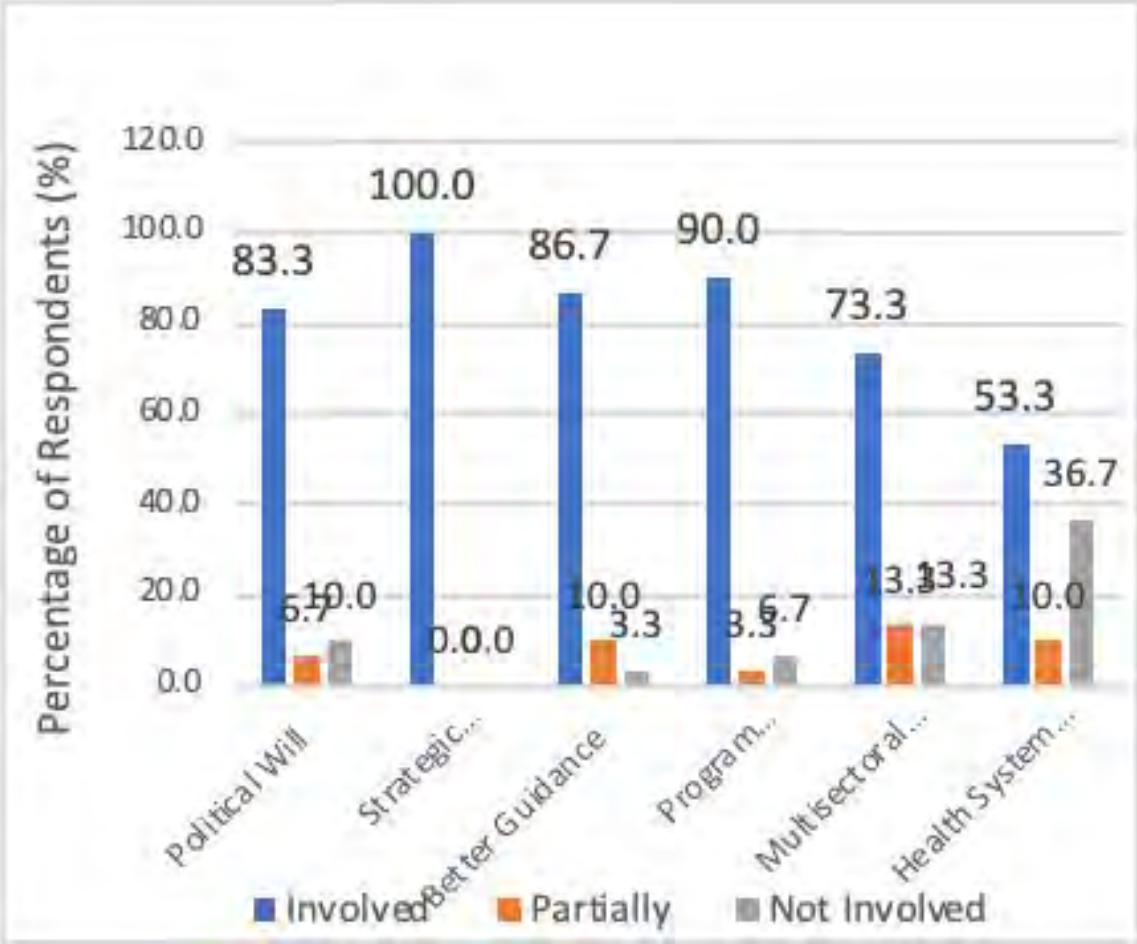
Country HBHI Initiation and Roll-out II

- Many of the health workforce engaged in the fight against malaria were not directly involved in the approach's launch, they learnt about it through interactions with National Malaria Programs (NMPs). NMP staff were actively engaged.
- Implementation challenges followed, with COVID-19 and resource issues hindering progress.
- At the subnational level, few knew the HBHI approach and its holistic implementation faced challenges.

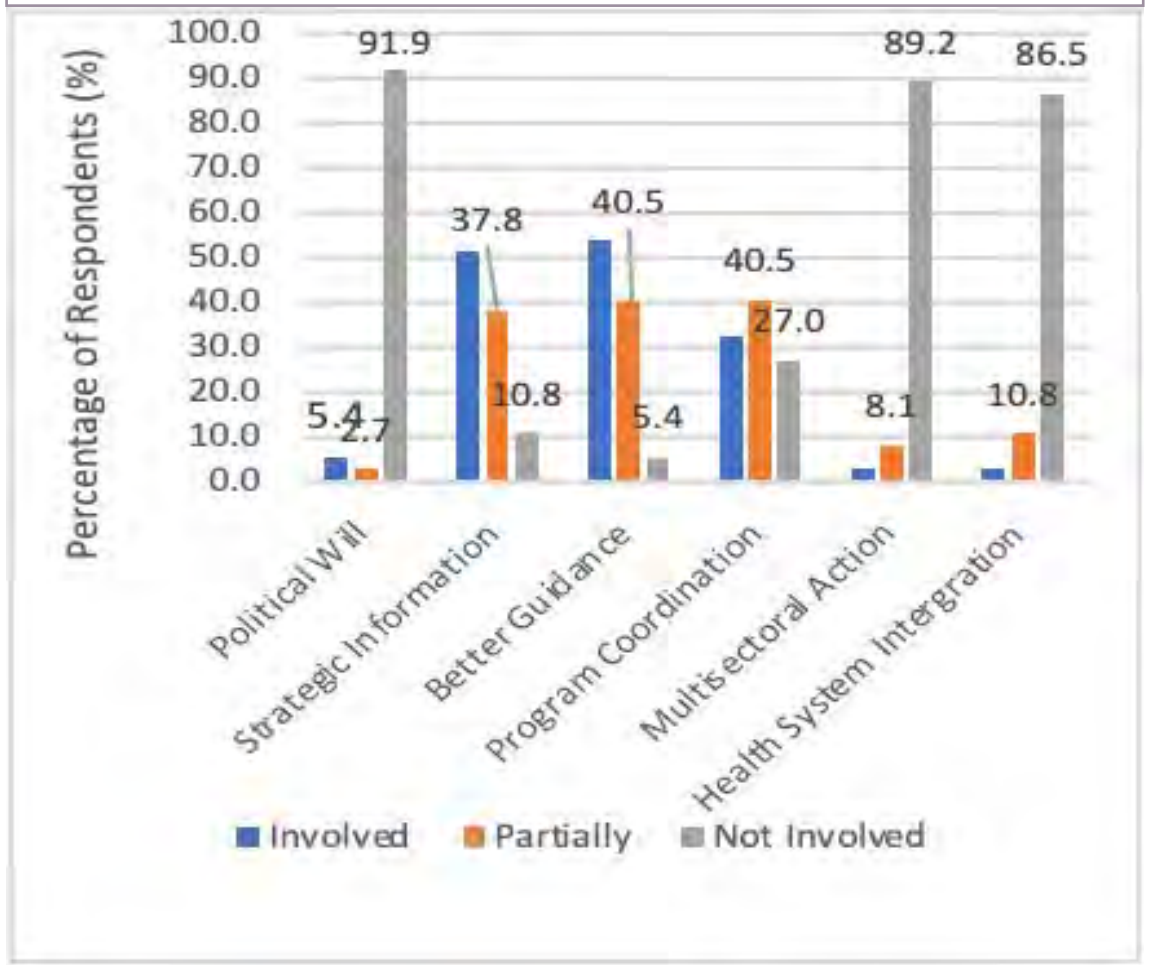
> *"Today, if you go into a health district and talk about HBHI, they'll think you're talking about something else. So, at the beginning, the enthusiasm was there, but after that, the management really wasn't what it should have been"*
> **(NMP Staff, Niger)**

Survey Respondents Involvement in HBHI

Ghana

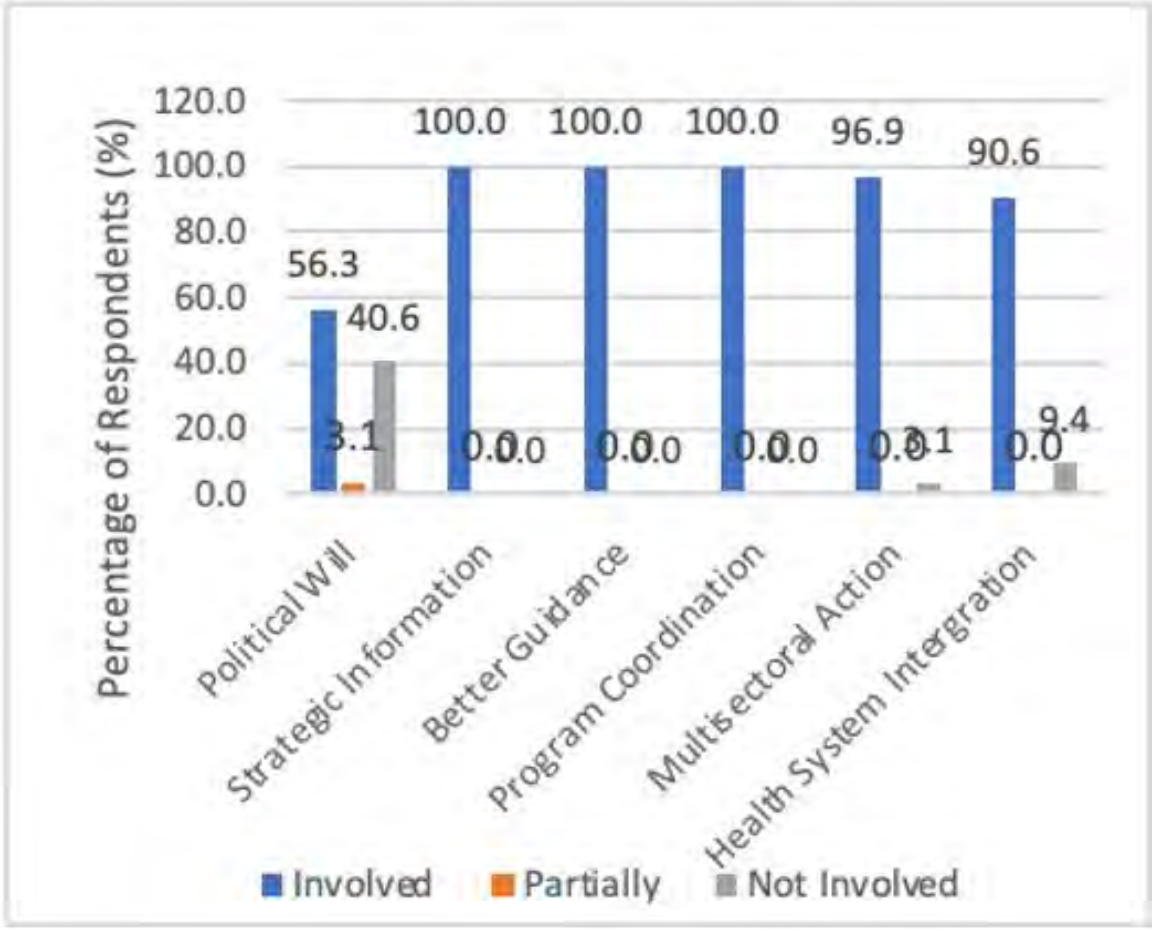


Mali

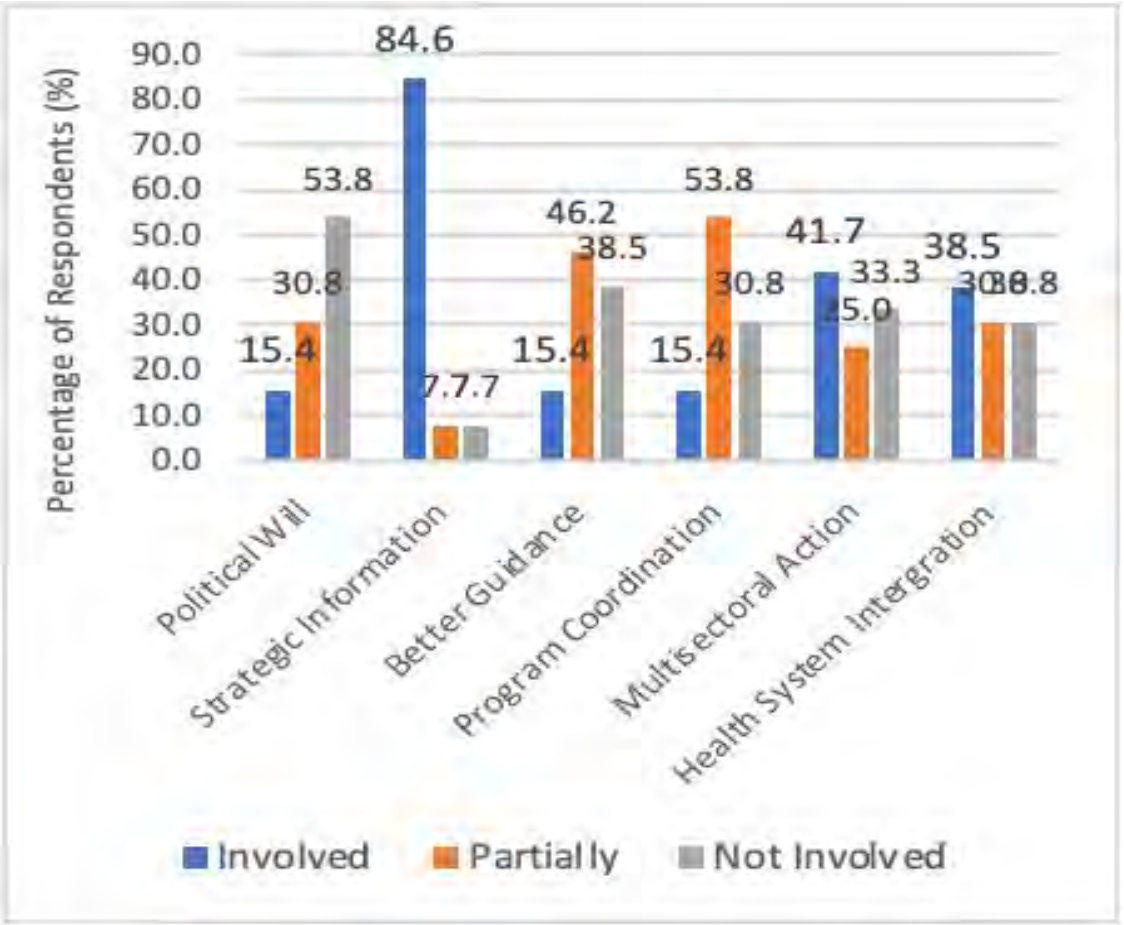


Survey Respondents Involvement in HBHI

Cameroun

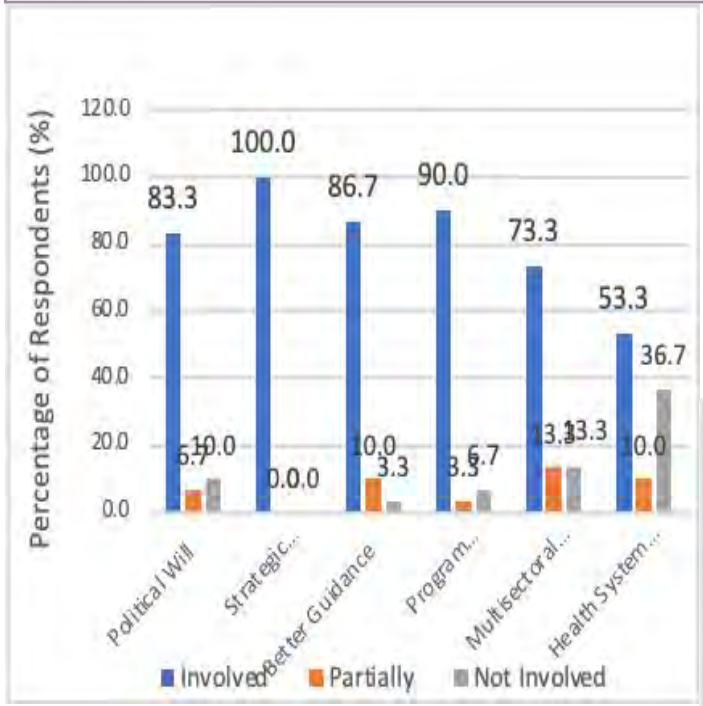


Niger

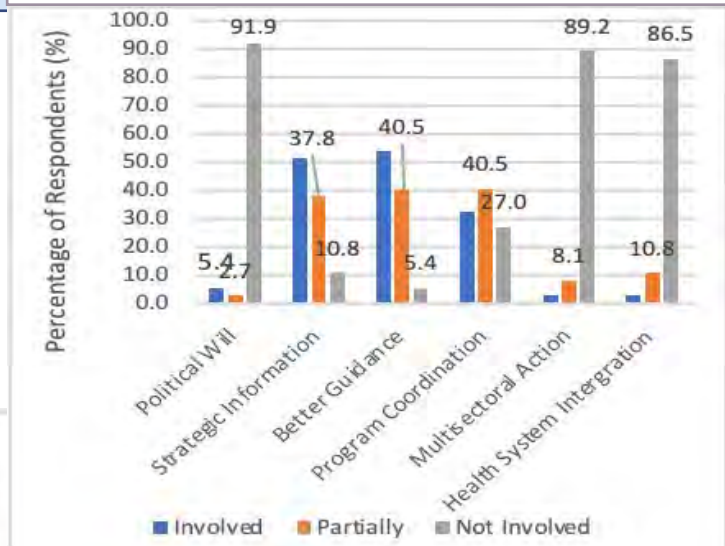


Survey Respondents Involvement in HBHI

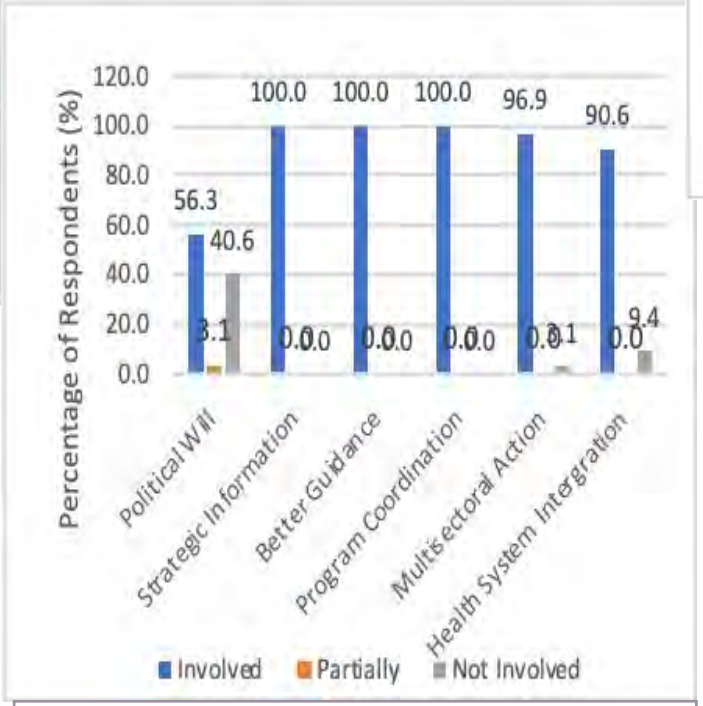
GHANA



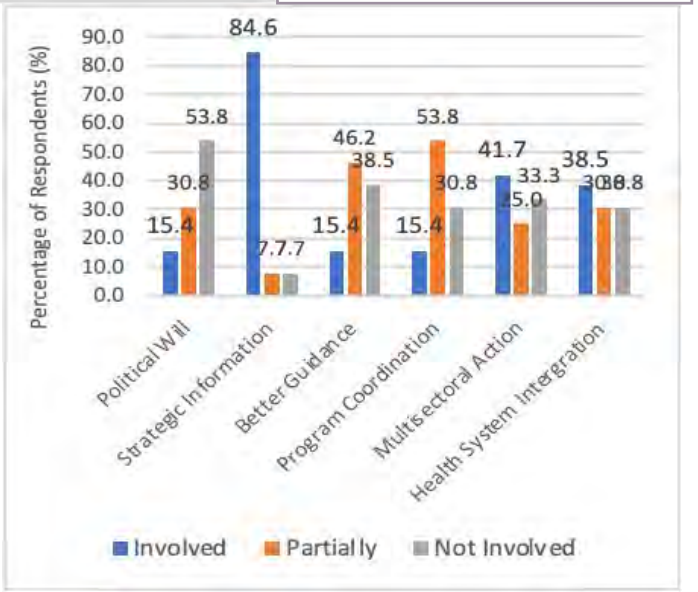
MALI



CAMEROUN



NIGER



World Health Organization

ership
Malaria

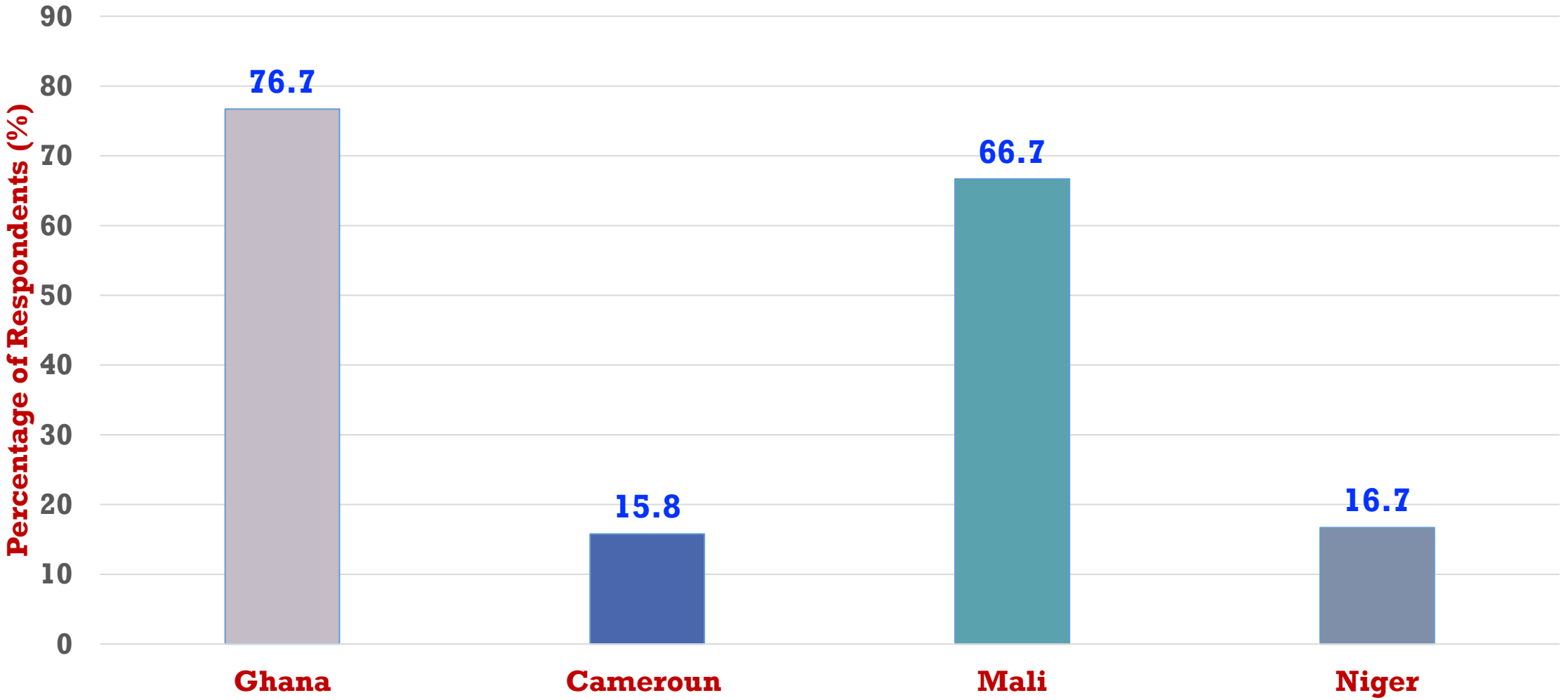
Influence of HBHI approach on malaria programme

- The HBHI approach has influenced malaria programs with varying degrees, particularly in improving strategic information and political will positively influencing National Strategic Plans (NSPs), Malaria Program Reviews (MPR) and Global Fund applications.
- A shift of paradigm from one-size fits-all to tailoring and better targeting of interventions and optimization of resources for high impact
- The approach drove updating of guidelines and the introduction of new ones to align with global policy guidance.
- At the global level, HBHI countries received increased attention and engagement from international stakeholders, resulting in more resources, visibility, and partnerships.
- **Subnational health managers and stakeholders generally lacked awareness of the HBHI approach and its implementation. Most had never heard of "HBHI"**

Pillar 1 – Political Will

SUCCESSES	CHALLENGES
Institutional changes and increased interest from political authorities and parliamentarians in malaria, particularly when supported by data.	Perception of politicians and local government officials that all health matters should be the responsibility of the health sector exclusively.
Creation of budget lines for malaria in some national budgets	Political changes and conflicts in some of the HBHI countries
Administrative upgrading of Malaria Programme to a National Directorate, demonstrating a strong national commitment to the fight against malaria	Slow change and little translation of political will to domestic resources
	General feeling "*malaria is with us; we just have to live with it.*"
	Competing priorities for limited resources
	Data or evidence that could encourage political leaders to invest in the fight against malaria are often not provided nor readily available to them
	End Malaria Councils not Functional

Perception of existence of adequate political structures in-country to ensure support for malaria



World Health Organization

RBM Partnership To End Malaria

Pillar 1 – Political Will

*"**The political will was expressed, but it was insufficient from an operational point of view.** The Ministry of Public Health should make a communication on the evolution of the malaria situation to all the Councils of Ministers. Malaria should then be considered a health priority in Niger, given that it is the leading cause of death in the country"* **(NGO, Niger)**

*"I'll say political will is one of the somewhat newer things the program is now focused on because of HBHI. **The HBHI initiative got us thinking outside the box of our regular unit to unit work and see how we can improve advocacy;** see how we can improve the visibility of malaria in Ghana so that we can move towards elimination"* **(NMP Staff, Ghana)**

Pillar 2 - Use of Strategic Information for Action

SUCCESSES	CHALLENGES
Marked Improvement in the use of Strategic Information for decision-making e.g development of malaria strategic plan, Sub national tailoring of intervention mix all based on most current data	Inadequate capacity for data analysis especially at sub national level limits continuity and effectiveness of data use.
The use of data and stratification has allowed for better decision-making, enabling countries to target interventions on areas with the highest malaria burden.	Lack of resources for data entry and management were, and equipment.
Integration of malaria data into national HMIS in some countries where this was not the case before HBHI	Poor data quality in the routine HMIS system and from private health facilities
	Annual reviews not resulting in any tangible changes.
	Repositories not yet functional in all 4 countries though they have all been initiated

Pillar 2 - Use of Strategic Information for Action

"Because we are an implementing partner and the funding that we have in our implementation, we could not cover the whole country, so that's why we adopted the HBHI approach. It was for us to prioritize where the need is, and we can implement our activities depending on the kind of intervention we are implementing that will inform us of the kind of indicators and data we need for prioritizing these areas. The other aspect is that most of the time, we use data a lot". **(Implementing Partner, Ghana)**

"It is stratification of all these interventions, which has made it possible to address the burden of the disease in fact, which has made it possible to implement the intervention packages in the target zones on the basis of this stratification. Personally, I think it was a great success in terms of stratifying the interventions and even implementing them." ***(NMEP Staff, Cameroun)***

Pillar 3 - Better Technical and Policy Guidance

SUCCESSES	CHALLENGES
Influenced to regular updating of guidelines in line with global WHO recommendations.	High personnel turnover affecting maintaining guideline adherence.
NMPs successfully disseminated updated guidelines to subnational levels.	Limited dissemination and reach (often during workshops) and leaving behind the majority of the health workforce.
Increasing collaboration with private sector stakeholders in guideline development and dissemination.	Poor accessing the latest guidelines, exacerbated by poor network connectivity.
Increased involvement of subnational levels on development, dissemination and optimal use.	Limited availability of hard copies and failure to share guidelines with colleagues.
	Resistance to adopting to new guidelines and practices.

Pillar 4: Programme Coordination

SUCCESSES	CHALLENGES
Functioning structures following elevation of malaria programme within the ministries of health	Partners often have their own targets and objectives and will not change them even if they are not aligned with those of the malaria programme
Increase in aligned partner support and harmonization	Lack of available domestic financial resources limits the ability of NMPs particularly at district level, to coordinate and ensure partners support adheres to national priorities.
Effective guidance and clarity on targets, objectives, outcomes and impact.	Re-designation of National Professional Officers (NPO's) to oversee several other diseases in addition to malaria is seen as a potential weakening of critical support by WHO to NMPs in their coordination effort.
Set up of Thematic Technical Working Groups (TWGs) comprised of by NMPs found to be a forum for improving programme coordination	

Pillar 4: Programme Coordination

"The partners in countries are engaged because that is why they are there in the first instance…… However, you need that bigger country ownership to direct how things are done and ensure that there's no duplication, things are maximized and gaps are filled appropriately. And one of the strengths of WHO in country presence is that they provide that backing for malaria programs to "speak up"." **(Global Respondent)**

"You know, the partners do everything they can to talk to each other, but it's not effective when the programme does not respond…… we wanted to put in place a mechanism to ensure that regular meetings were held at least once a quarter….it's a pity when there isn't regular dialogue between those who fund and those who receive the funding in the implementation… and that's the challenge….. **(Development Partner, Niger)**

Multisectoral Action

SUCCESSES	CHALLENGES
Some sectors, such as education have had and continue to have a well established collaboration with the NMPs.	Difficulty of attaining leadership from highest levels of national authorities (presidency or prime ministry), to ensure all sectors collaborate or link in planning, financing, implementing and accountability measures.
Clear understanding of how various sectors can contribute to malaria through one health	The NMPs did not feel they had the needed leverage to bring other sectors outside of the health together, in view of their position in the hierarchy of the health service.
	Lack of understanding of the malaria problem by other sectors
	Little progress in implementation despite national plans or frameworks
	Inadequate resources and unclear ownership and stewardship

Health System Integration and Covid-19

SUCCESSES	CHALLENGES
On-site training and supportive supervision (OTSS) involving Malaria, TB, HIV and Maternal, Newborn and Child Health	Stakeholder engagement and buy-in for HBHI were hindered by the absence of interpersonal engagement due to the pandemic.
The COVID-19 response provided valuable lessons for malaria in HBHI countries.	Resources diversion to COVID-19 response, affecting all HBHI pillars
Integrated supervision activities of District Health Management Teams on Malaria/HIV/Maternal, Newborn and Child Health (MNCH) provides an opportunity to validate data from the Community Heath Centres (ComHCs)	Service disruption and postponement of malaria campaigns, resulting in increased delivery costs for interventions.
Incorporation of malaria control into the iCCM package for community health workers in one country	
Initial engagement with political circles as part of Political Will laid the foundation for the malaria program representation within the covid mitigation committee.	

Lessons from Global Implementation Processes

- Immense support was received by countries from RBM and WHO in preparatory activities towards stakeholder engagement, high level political engagement and launch of the HBHI approach.
- There were processes in place to provide technical assistance with regards to some of the pillars.
- Beyond log frames developed by in-country stakeholders to guide implementation, no guidelines were provided
- Additionally, there was no monitoring and & evaluation framework embedded into the implementation.
- Coordination meetings were not as effective as they should have been. They did not involve the NMPs and frequency of the meetings did not allow time for any concrete changes to be made
- Challenges still remain, in translating the political will generated into increased domestic resources and reflecting the central will generated at sub national levels.

General Lessons

- Perspective of NMPs on malaria control /elimination were broadened
- HBHI was generally perceived as a project and not an approach
- There was lack of ownership of the HBHI approach by some program staff and national level stakeholders
- There was an expectation of additional funding to support implementation which did not materialize leading to decreased momentum and enthusiasm over time
- The lack of a forum for peer learning among the country implementers (NMP staff and their stakeholders) even if virtual was a missed opportunity

General and Global lessons

"I'll say that flagging these countries as high burden and top priority has really catalyzed the attention from the global fund, from PMI, from global stakeholders on the importance of these countries….." **(Global Respondent)**

"The HBHI approach has not concretely brought new principles, but new orientations have really been adopted by the programme in relation to HBHI approach. I think that through this, the programme has directed its efforts…." **(NMP Staff, Mali)**

"They all succeeded at the initiation of the HBHI, they all came up with and identified the gaps and the plans. Those plans were never really systematically implemented and all we kept doing was bringing them together, they came repeating to us the plans as part of the presentation." **(Global Respondent)**

RECOMMENDATIONS

Recommendations

- **Recommendation 1- Ensure a mind shift from viewing HBHI as a project to seeing it as an approach.**
- **Recommendation 2 - Rethink how to move from expressed political will to a translation to increased domestic resources**
- **Recommendation 3 - Tailor HBHI Efforts to each country's specific context and strengths.** While some elements may be easier to address in certain countries, the overall success of the approach hinges on a comprehensive approach that aligns with the unique needs and capacities of each participating country.
- **Recommendation 4 - Build capacity of Malaria Programme managers and Health Managers to lead the effort.** This requires a change in the way the malaria programme is traditionally run and also a move from comfort zones to areas that are not the norm for programme staff, their stakeholders and political authorities. Of particular importance is an expansion of the skills set to include softer skills such as social and behavioural sciences and political analytics.

Recommendations II

- **Recommendation 5 - Additional human resource allocation and funding is vital:** Human resource allocation and some specific funding is vital for efficient execution and coordination of the approach. Some of the areas of HBHI activities are not included in country budgets traditionally and this needs to be reviewed. Additional human resource may be needed in the short to medium term in some cases.

- **Recommendation 6 - Decentralize HBHI and ensure that it is deployed at all other sub national levels:** While high-level political involvement is deemed essential at the national level, it is important to find ways to decentralize the HBHI approach down to all levels in order to maximize the impact of the approach. This has to be intentional and requires funding to accomplish. This will then mainstream the thinking and approach in what all the different levels are doing to eliminate malaria.

Recommendations III

- **Recommendation 7 - Set up a forum for sharing experiences and lessons among implementing countries.** It should be country-led but facilitated and supported by one of the global partners
- **Recommendation 8 - Develop Implementation Guidelines to provide guidance to countries** based on the experiences gathered from the roll-out in the initial set of countries
- **Recommendation 9 - Build in a Monitoring and Evaluation Framework into the Approach.** A monitoring and evaluation framework needs to be built into the implementation of the HBHI approach upfront so that this is done at regular intervals with the results feeding into an improvement in the implementation and course correction where necessary
- **Recommendation 10 – Effective Multisectoral action will ride on political will at the highest level and should be a key outcome of the Political will pillar.**

CONCLUSION

Conclusion

- The implementation of the High Burden to High Impact (HBHI) approach has had varying degrees of success in improving programme performance at country level.
- Notably, there have been significant advancements in certain areas, particularly in strategic information and political will. In these areas, time and effort have yielded obvious improvements.
- The extent of progress has however been slightly constrained by challenges related to coordination, the development of better guidance and policies locally, ability to effectively engage other sectors and critical non-health stakeholders in the fight against malaria.
- These areas have shown comparatively less progress. Moving beyond expressed political will to increased domestic financing has also proved to be a bottleneck.
- Roll-out of the HBHI approach has still not moved beyond the national level with a general lack of awareness below the national level.

Wholistic approaches are the solution to fight malaria in Africa

Emina J, Yé Y, PLOS ONE 16(5), 2021

Thank you

Update on WHO malaria guidelines upcoming reviews on primaquine, tafenoquine and G6PD near patient diagnostic tests

Dr Andrea Bosman and Dr Peter Olumese, GMP Diagnostic, Medicines and Resistance Unit

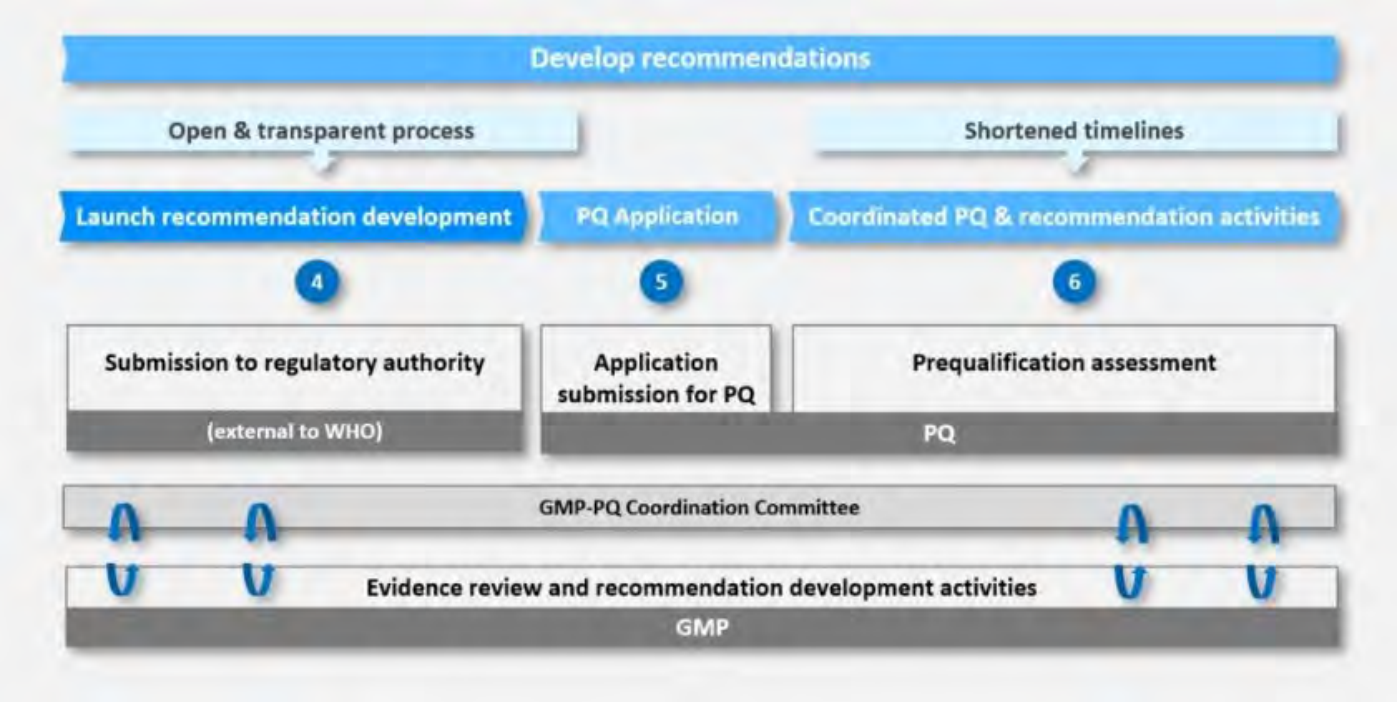
24th meeting of the WHO Malaria Policy Advisory Group

30 October – 1 November 2023

World Health Organization

Background

- Based on the review of development of WHO malaria recommendations, in Nov 2019 GMP and the Department of Essential Medicines Health Products developed and Master Plan for Developing Recommendations on the Use of Tafenoquine and companion Quantitative Point-of-Care G6PD In Vitro Diagnostic(s)
- Aim to coordinate activities as 'one WHO' on recommendations on the use of tafenoquine and companion G6PD POCT as part of:
  - WHO Guidelines for Malaria,
  - WHO prequalification list of prequalified finished pharmaceutical products and in-vitro diagnostics
  - Model List of essential medicines and essential diagnostics.



https://www.who.int/teams/global-malaria-programme/guideline-development-process/recommendation-pathway

World Health Organization

Timelines for WHO recommendations on tafenoquine and primaquine

14-15 Nov 2023

GDG meeting: review evidence, formulation of recommendations

Systematic reviews on
- **Tafenoquine**
- **Primaquine**

Feasibility study (Brazil)

Nov - Dec 2023

Recommendations drafted and approved by GDG

Jan 2024

External review group provide peer review

Feb - Mar 2024

Editing, GRC internal review

April 2024

Publication of recommendations on the use of tafenoquine and primaquine, if positive guidelines recommendations and PQ listing of quantitative near-patient G6PD tests

World Health Organization

PICO question on tafenoquine anti-relapse therapy

Topic	PICO Question	Inclusion criteria				Critical & important Outcomes	Current status
		Studies	Participants	Intervention	Control		
Single-dose tafenoquine for radical cure of *Plasmodium vivax* malaria (to be recommended with a near-patient quantitative G6PD test)	Is single dose tafenoquine an alternative to standard dose primaquine for preventing relapses in patients with a G6PD activity of >70% who have received chloroquine therapy for acute *P. vivax* infection?	All eligible studies Including Phase IV studies for safety review	Patients with a G6PD activity of >70% treated for *P. vivax* malaria with chloroquine	Single dose tafenoquine (300mg)	Standard Primaquine treatment 0.25mg/kg daily for 14days or 0.5mg/kg daily for 7 days or 0.5mg/kg daily for 14 days or placebo	*P. vivax* relapse defined as reappearance of *P. vivax* parasitemia <6 months after treatment Safety of tafenoquine	New evidence review

World Health Organization

PICO question on primaquine anti-relapse therapy

Topic	PICO Question	Inclusion criteria				Critical & important Outcomes	Current status	Other considerations/ questions
		Studies	Participants	Intervention	Control			
Anti relapse treatment: Primaquine efficacy	Is high total dose primaquine (7.0 mg/kg) more efficacious than low total dose primaquine (3.5 mg/kg) at preventing relapses to day 180 in patients with uncomplicated vivax malaria?	Clinical trials and cohort studies	Malaria patients with *P. vivax* uncomplicated disease	PQ at high total dose (7.0 mg/kg)	PQ at standard/ low total dose (3.5 mg/kg)	First vivax recurrence by day 180	Update evidence review	Impact of duration of treatment Impact by geographic region Impact of schizontocidal drug Age <5 years
Anti-relapse treatment: Primaquine – tolerability and safety	Does intermediate (0.5 mg/kg) or high (1.0 mg/kg) daily dose primaquine cause more gastrointestinal symptoms or adverse haemoglobin changes compared to low (0.25 mg/kg) daily dose primaquine?	Clinical trials and prospective cohort studies.	Malaria patients in vivax endemic regions with uncomplicated disease and G6PD activity >=30%.	Daily PQ dose of 0.5 mg/kg Daily primaquine dose 1.0 mg/kg	Daily PQ of 0.25 mg/kg	Vomiting or diarrhoea or anorexia on Days 2-3 and 5-7 Vomiting within 1st hour Hb change on Days 2-3 Hb drop >25% to <7 g/dL on Days 1-13	Update evidence review	G6PD activity: >=30%, 30-<70% and >=70% Sex Food intake GI symptoms in relation to age

PICO question on primaquine

Topic	PICO Question	Inclusion criteria				Critical & important Outcomes	Current status	Other considerations/ questions
		Studies	Participants	Intervention	Control			
Primaquine for infants aged < 6 months and breastfeeding women	Is it safe to administer primaquine to infants aged < 6 months and women breastfeeding infants aged < 6 months to reduce transmission and to prevent relapses?	Safety surveillance Case reports	Children and lactating women with uncomplicated *P. falciparum* or *P. vivax* malaria	Infants <6 months Women breastfeeding infants aged <6 months	Infants >6 months Women breastfeeding infants aged >6 months	Safety (serious adverse events, haemolysis and vomiting) Drug levels in breastmilk	New evidence review	G6PD status
Single low dose primaquine for reducing spread of artemisinin resistance	In areas threatened by artemisinin resistance, a single low dose of primaquine of 0.25 mg/kg should be given with ACT to patients with *P. falciparum* malaria	RCTs Cohort studies Observational studies	*P. falciparum* malaria infected patients	In moderate to high transmission intensity areas	In low transmission areas	Gametocyte carriage (qPCR, microscopy) Mosquito membrane feeding experiments (using ex vivo blood samples)	Update evidence based on IPD meta-analysis (Efficacy of SLD PQ • JID 2022:225 (1 April)	Age Seasonality of malaria Pre-treatment: - Parasite density - Gametocytaemia - Hb - Duration of illness ACT PQ dose

Timelines for WHO recommendations on near-patient G6PD tests

30 Nov – 1 Dec 2023

GDG meeting: review evidence, formulation of recommendations

Dec – Jan 2023

Recommendations drafted and approved by GDG

Feb 2024

External review group provide peer review

Mar 2024

Editing, GRC internal review

April 2024

MAGIC Evidence Ecosystem Foundation

Publication of recommendations, if PQ listing of G6PD near patient semi-quantitative tests

Systematic reviews on

- **Qualitative G6PD point-of-care tests**
- **Quantitative G6PD point-of-care tests**

Literature review of contextual factors (feasibility, acceptability, gender and equity)

Linked evidence modelling of the impact of G6PD testing strategies on treatment outcomes of *P. vivax*

World Health Organization

PIRT question on near-patient G6PD tests

Topic	PICO Question	Inclusion criteria				Critical & important Outcomes	Current status	Other considerations/ questions
		Studies	Participants	Index test	Reference Standard			
Use of near-patient qualitative or quantitative G6PD tests to support safe and effective *P. vivax* and *P. ovale* anti-relapse treatment	In patients undergoing G6PD activity testing, how accurate are near-patient tests for G6PD deficiency compared to quantitative spectrophotometric G6PD testing at the thresholds* critical to inform administration of 8-aminoquinolines to prevent relapses of P. vivax and P. ovale? * <30% vs 30-70% vs >70% G6PD activity	All eligible studies	Patients undergoing G6PD testing	- G6PD FST - CareStart G6PD - BinaxNOW G6PD - WST8/1-methoxy PMS assay - Standard G6PD by SD Biosensor - CareStart G6PD Biosensor by AccessBio	Quantitative spectro-photometric assay The reference G6PD activity (100%) calculated as adjusted male median of study samples for each spectro-photometric assay	Sensitivity and specificity at 30% and 70% G6PD activity in males and females	New evidence review	Gender Age G6PD prevalence Endemicity of malaria Location (e.g. Africa, Asia) Venous vs capillary Reference standard

PIRT = Population, Index Test, Reference Test, Target Condition

Thank you
for your attention

World Health Organization

Update on status of antimalarial drug resistance in Africa

Charlotte Rasmussen

Global Malaria Programme

World Health Organization

Outline of presentation

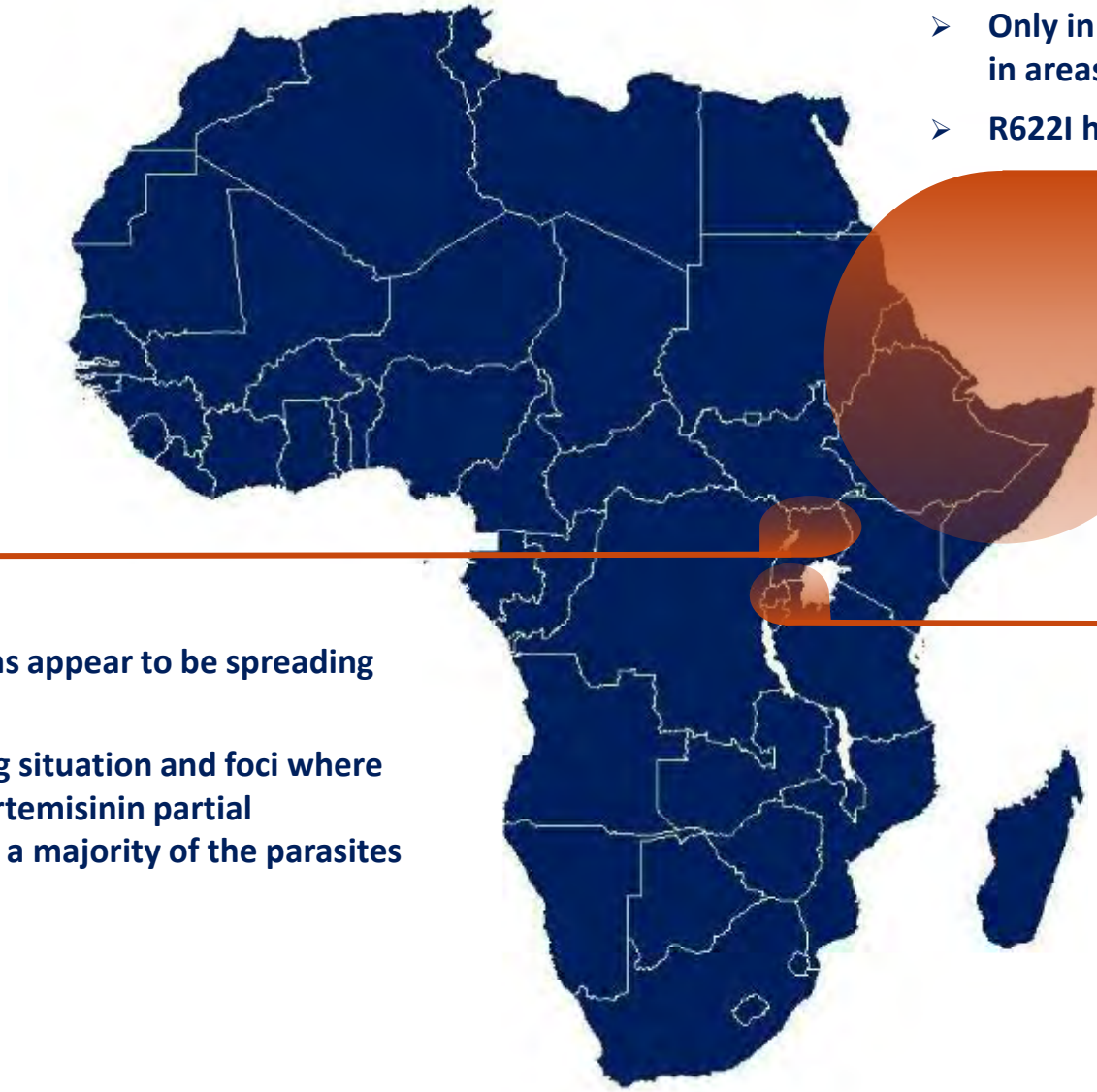
- Background
- Current information on antimalarial drug resistance
- Updates on WHO activities to operationalize the strategy to respond to antimalarial drug resistance in Africa

World Health Organization

Background

- The **Strategy to respond to antimalarial drug resistance in Africa** was launched in November 2022 following an extensive process that included an MPAG review
- A review done as part of the development of the strategy found that **artemisinin partial resistance at that time had been identified in 3 countries** in Africa:
  - For **artemisinin partial resistance to be confirmed in a site**, quality evidence is needed on:
    - ✓ Presence of validated marker (≥5%) (*PfK13* mutations)
    - ✓ Evidence of delayed clearance (Day 3 + or parasites clearance half-life)
- The review also found that there were scattered reports of high treatment failure rates but **no confirmed ACT partner drug resistance**

Current pattern of artemisinin partial resistance



Horn of Africa

- K13 mutation R622I detected in several countries in the Horn of Africa including Eritrea, Ethiopia, Sudan and Somalia
- Only in Eritrea is there evidence of delayed parasite clearance in areas of high prevalence of R622I
- R622I has been detected in parasites with *Pfhrp2/3* deletions

Uganda

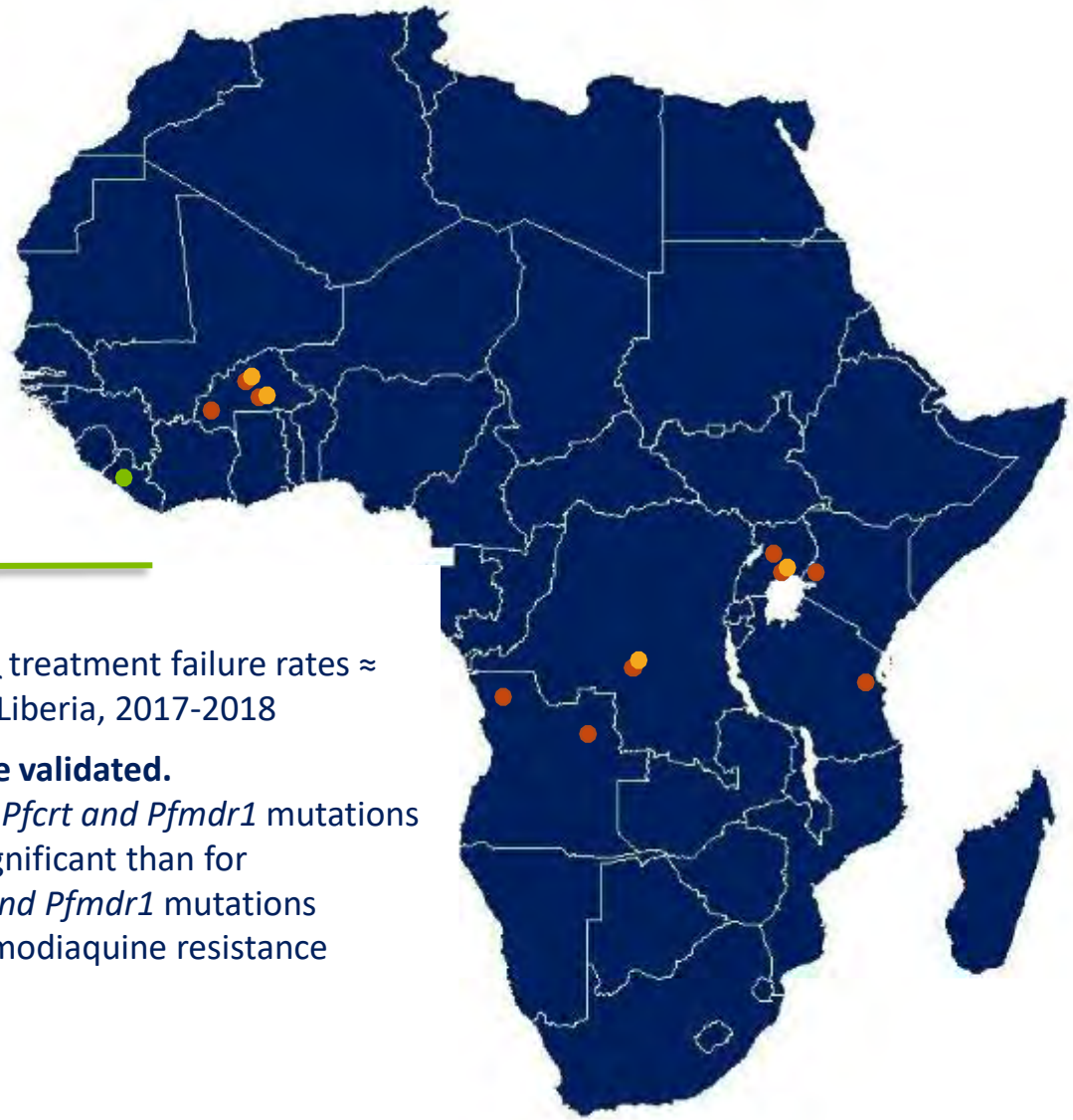
- Different K13 mutations appear to be spreading in Uganda
- Data shows an evolving situation and foci where validated markers of artemisinin partial resistance are found in a majority of the parasites sampled

Rwanda & Tanzania

- K13 mutation R561H had been found at high prevalence in studies with evidence of delayed clearance in Rwanda
- R561H has now also been detected in Tanzania in a study with a high proportion of patients with delayed clearance indicating the presence of artemisinin partial resistance in Tanzania

World Health Organization

Conclusion in 2022 review: Scattered reports of high treatment failure but no confirmed partner drug resistance in Africa



TES with high failure rates from 2015 - 2023

Amodiaquine

- **Current evidence:** ASAQ treatment failure rates ≈ 10% identified in TES in Liberia, 2017-2018
- **Molecular marker: To be validated.**
  IC50 affected in vitro by *Pfcrt and Pfmdr1* mutations but shift of IC50s less significant than for chloroquine, and *Pfcrt and Pfmdr1* mutations cannot be considered amodiaquine resistance markers at present

Lumefantrine

- **Current evidence:** AL treatment failure rates > 10% reported in Angola, Burkina Faso, Democratic Republic of Congo, Kenya, **Tanzania and Uganda**
- Increased IC_{50} in Uganda
- **Molecular marker: To be validated**
  Studies show that lumefantrine selects for *Pfmdr1* N86

Comments: Different challenges with TES for AL

- Short half-life -> potential misclassification of reinfections as recrudescences
- Some studies have used PCR-correction method based a Bayesian algorithm, some concerns on quality of microscopy, and some studies without supervision of evening dose
- High reinfection rates in some sites

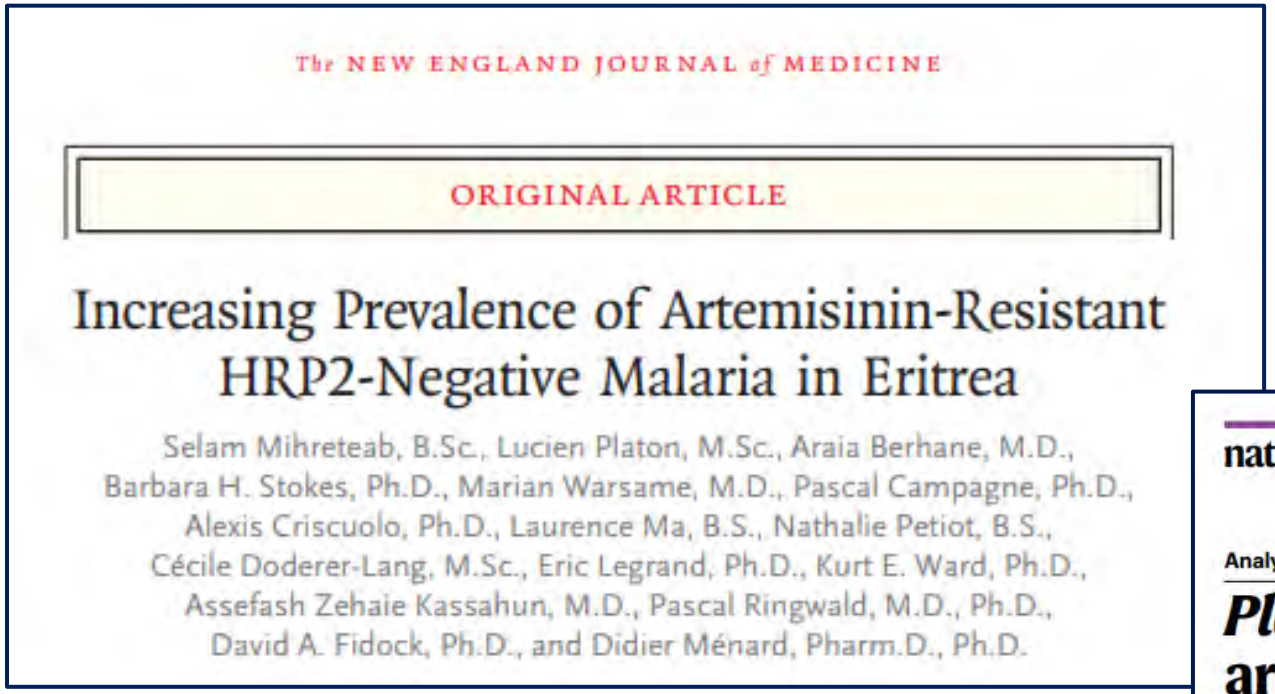
Piperaquine

- **Current evidence:** DP treatment failure rates > 10% or ≈ 10% reported in Burkina Faso, the Democratic Republic of Congo and Uganda
- **Molecular marker:** *Pfpm2–3* increased copy number and *Pfcrt* mutations validated in Asia and South America

Comments:

- Studies have used PCR-correction method a Bayesian algorithm & some concerns on quality of microscopy
- In Burkina Faso, Uganda and DR Congo, DP treatment failures in sites where AL treatment failures were also found in studies using Bayesian algorithms for PCR corrections

Artemisinin partial resistance - Horn of Africa

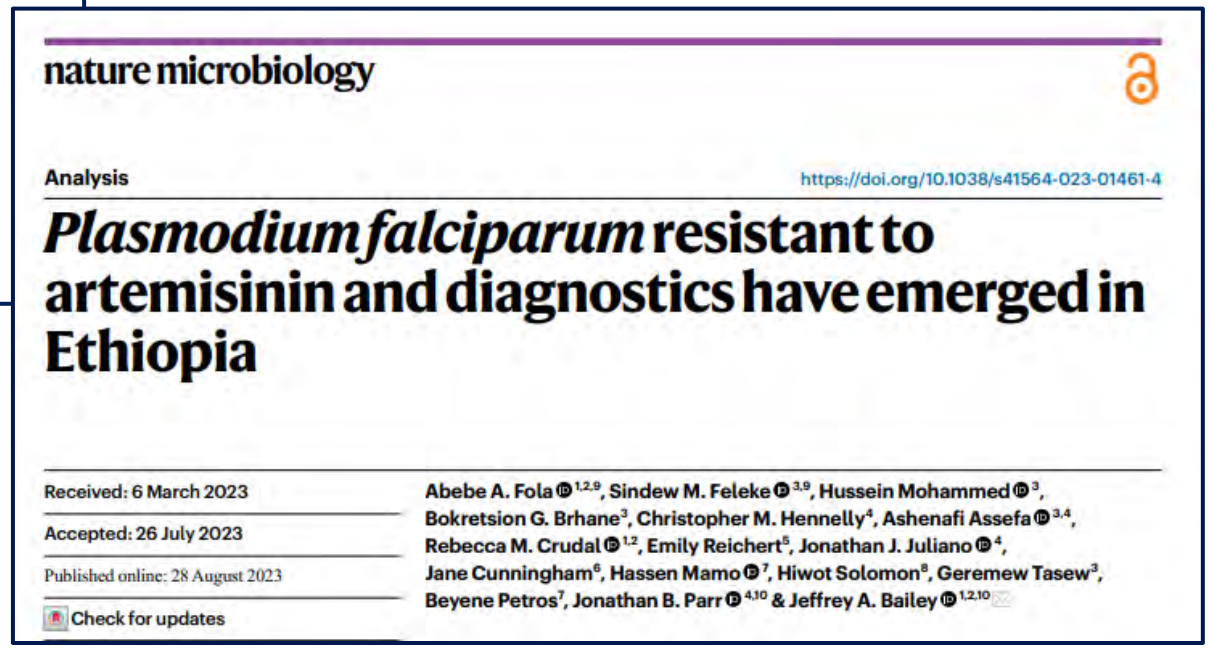


The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Increasing Prevalence of Artemisinin-Resistant HRP2-Negative Malaria in Eritrea

Selam Mihreteab, B.Sc., Lucien Platon, M.Sc., Araia Berhane, M.D., Barbara H. Stokes, Ph.D., Marian Warsame, M.D., Pascal Campagne, Ph.D., Alexis Criscuolo, Ph.D., Laurence Ma, B.S., Nathalie Petiot, B.S., Cécile Doderer-Lang, M.Sc., Eric Legrand, Ph.D., Kurt E. Ward, Ph.D., Assefash Zehaie Kassahun, M.D., Pascal Ringwald, M.D., Ph.D., David A. Fidock, Ph.D., and Didier Ménard, Pharm.D., Ph.D.



nature microbiology

Analysis

https://doi.org/10.1038/s41564-023-01461-4

Plasmodium falciparum resistant to artemisinin and diagnostics have emerged in Ethiopia

Received: 6 March 2023

Accepted: 26 July 2023

Published online: 28 August 2023

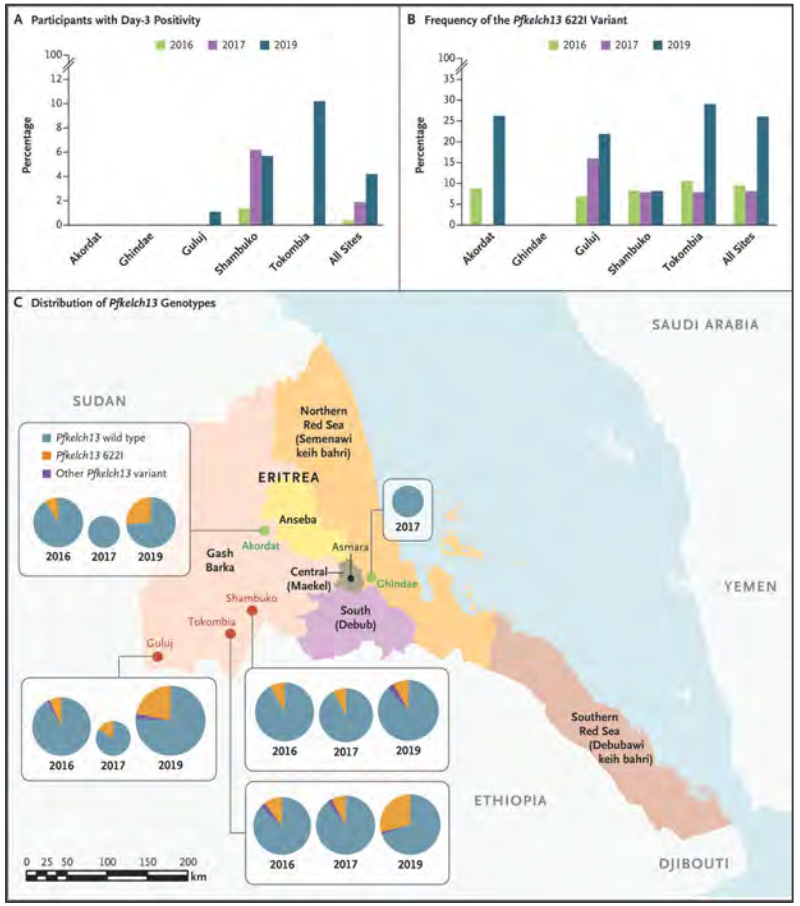
Check for updates

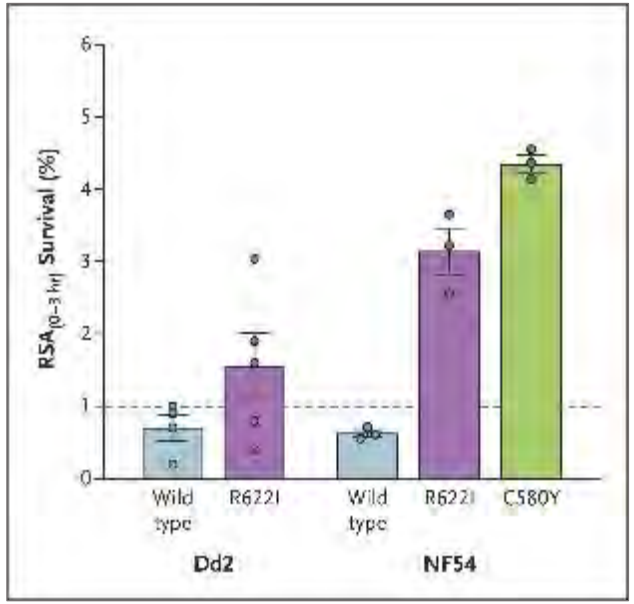
Abebe A. Fola[1,2,9], Sindew M. Feleke[3,9], Hussein Mohammed[3], Bokretsion G. Brhane[3], Christopher M. Hennelly[4], Ashenafi Assefa[3,4], Rebecca M. Crudal[1,2], Emily Reichert[5], Jonathan J. Juliano[4], Jane Cunningham[6], Hassen Mamo[7], Hiwot Solomon[8], Geremew Tasew[3], Beyene Petros[7], Jonathan B. Parr[4,10] & Jeffrey A. Bailey[1,2,10]

World Health Organization

Eritrea

- Published data from samples collected 2016 – 2019 shows an increased prevalence of *Pfkelch13* 622I & evidence of delayed clearance
- Ring stage surveillance assay shows that R622I mutation conferred a low level of resistance to artemisinin when edited into African (NF54) and Asian (Dd2) parasite lines
- Of parasites with *Pfkelch13* 622I, 16.9% had both *hrp2* & *hrp3* deletions. In wild-type parasites, 21.8% had both *hrp2* & *hrp3* deletions
- Unpublished TES data from 2022 show:
  - Very high efficacy of ASAQ
  - An increase in day 3+ in one site (Shamboko: 23.6% Day 3+)
  - Partial analysis of samples show 21% - 43% prevalence of R6622I
  - High *hrp2/3* deletions (50% with *hrp2* deletions and 43.5% dual deleted)





Mihreteab et al. N Engl J Med Sep 2023

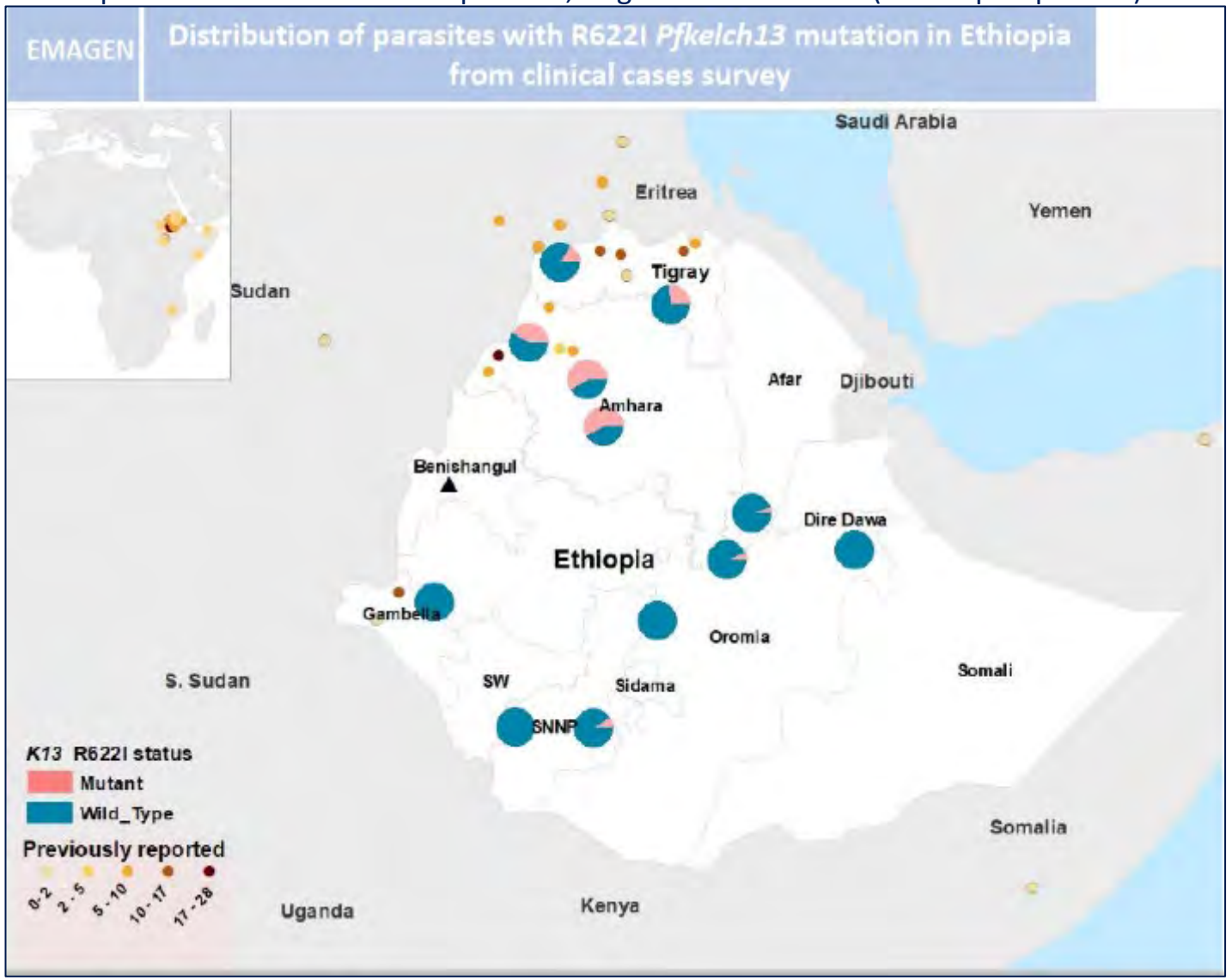
Ethiopia

- *PfKelch13* 622I highly prevalent particularly in northern Ethiopia
- *Hrp2/3* deletions common. Analysis of data from 2017 – 18, showed 622I mutation is more common among pfhrp2/3 non-deleted parasites (11.6%) than among pfhrp2/3 double-deleted parasites (4.5%). However, this varies between studies and sites
- 2022 TES data from one site in Amhara shows good efficacy (>90%) of artemether-lumefantrine and artesunate-pyronaridine, and very low day 3 positivity.

2022 TES data

Study arm (enrolled)	Treatment outcome		PCR-adjusted
	ACPR, %(n/N)	LPF, %(n/N)	ACPR, %(n/N)
28 days			
Pf: AL-PQ (101)	96.8 (91/94)	3.2 (3/94)	97.8 (90/92)
Pf: PA-PQ (98)	97.8 (88/90)	2.2 (2/90)	98.9 (88/89)
42 days			
Pf: AL-PQ (101)	94.3 (83/88)	5.7 (5/88)	96.5 (83/86)
Pf: PA-PQ (98)	93.1 (81/87)	6.9 (6/87)	94.2 (81/86)

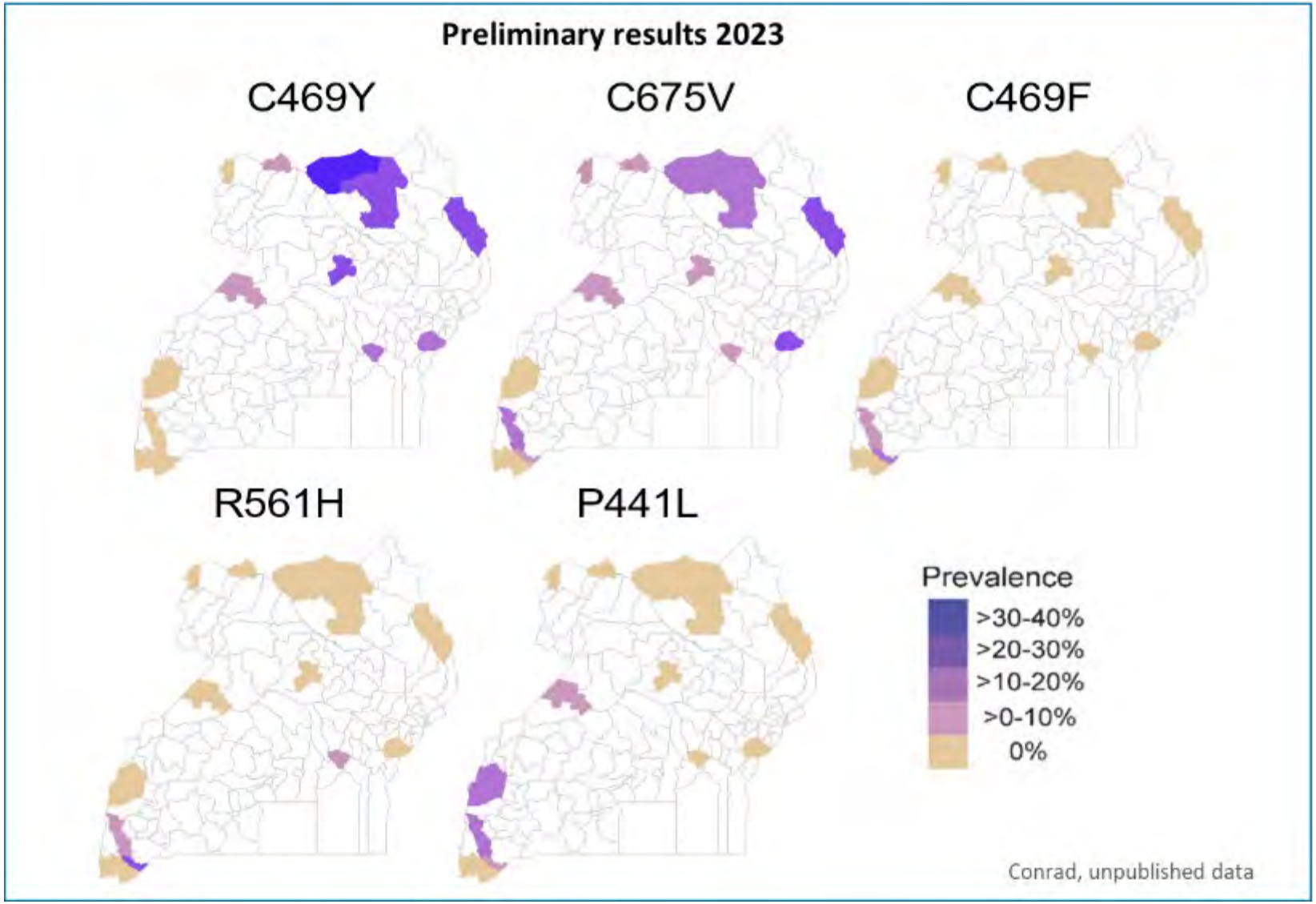
Samples collected from malaria patients, Aug. 2021 – Dec 2022 (50 samples per site)



Source: Presentation, F. Girma

Artemisinin partial resistance - Uganda

- In Uganda, data from 2016 – 2022, shows that different *PfK13* mutations are becoming more prevalent and spreading geographically
- Preliminary results from 2023, show some mutations such as C469Y and C675V highly prevalent in northern Uganda while other mutations including R561H is more prevalent in southern Uganda.
- Analyses suggest single origin on 469F and 469Y but potentially 2 origins of 675V and 561H (Uganda vs Rwanda)
- Analysis of the spread of the mutations in Uganda indicate that the selection coefficient is equivalent what was seen in the Greater Mekong subregion*



*Meier-Scherling, unpublished

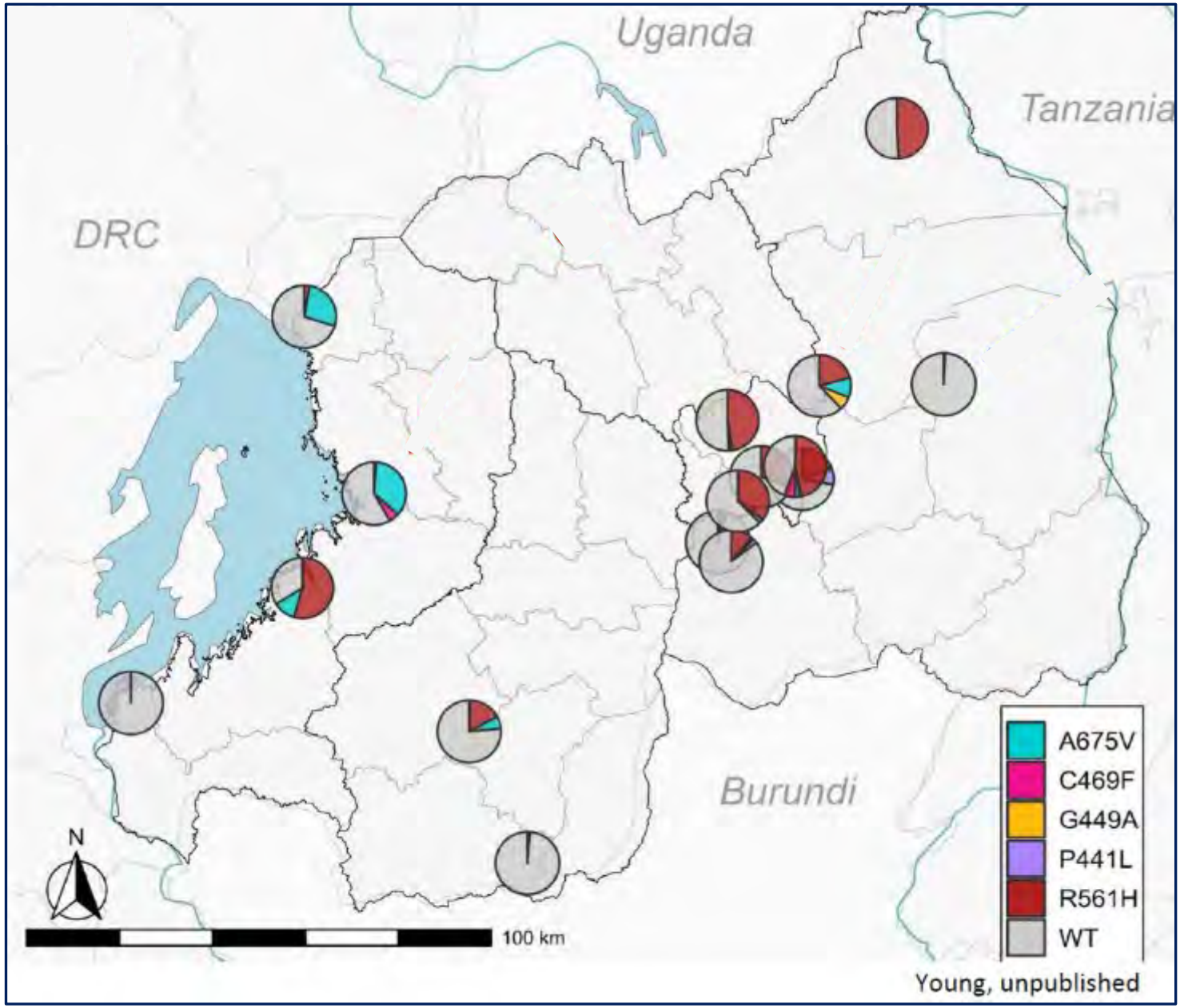
Rwanda

Genomic surveillance

- Genomic surveillance was done Feb 2022 – March 2023, analyzing 2713 samples (in 104 pools) from 21 locations.
- The analysis found that K13 561H spreading rapidly in Rwanda
- A675V more common in western Rwanda

In-vivo data

- Delayed parasite clearance at frequencies ranging from 14% to 28% has been detected in 3 sites in 2018 and 2019. AL and DHA-PQ failure rates under the 10% threshold.
- PCR corrected data from the 2022 TES (3 sites) are not yet available. Uncorrected results indicate a stable situation



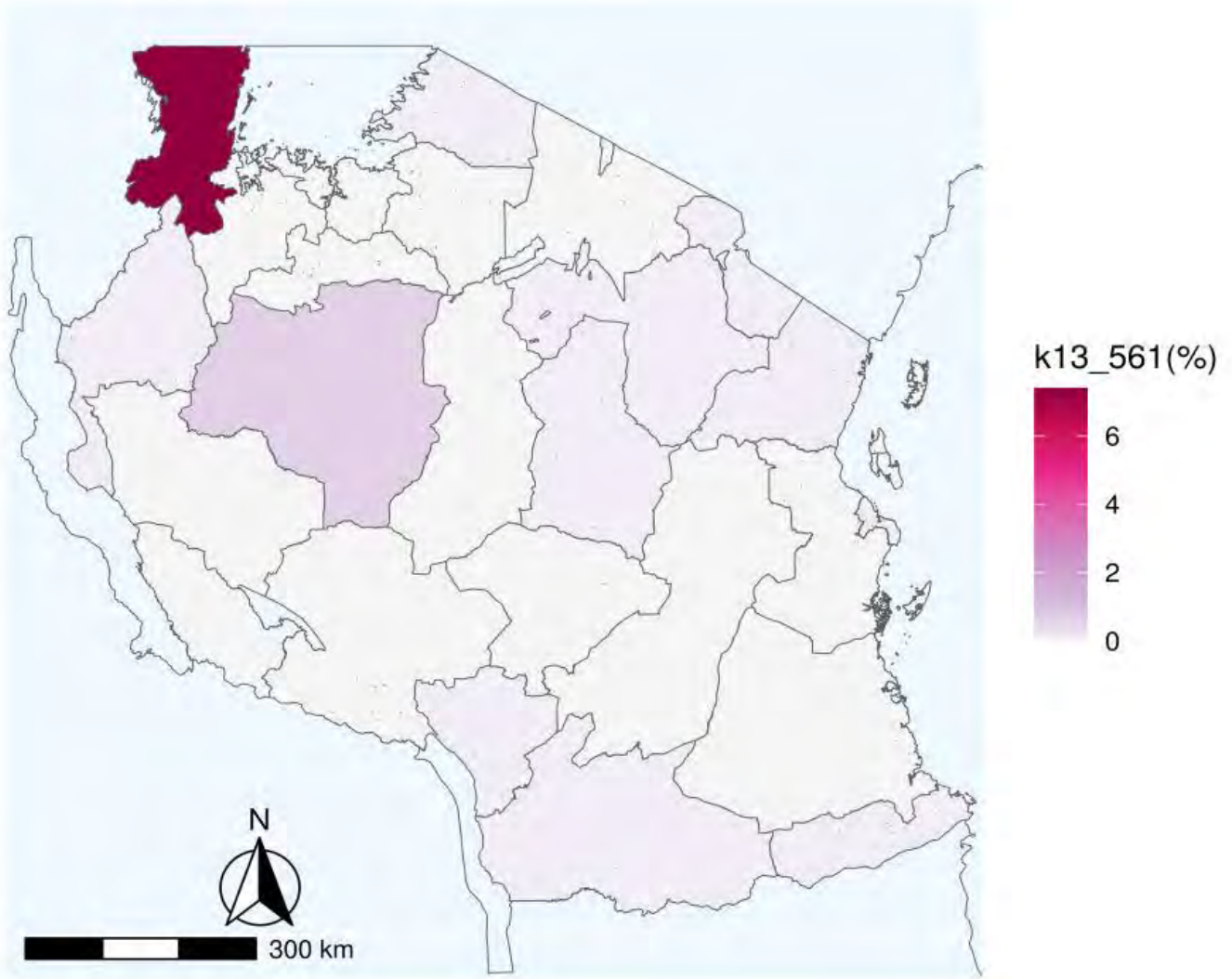
Source: Jeff Bailey, Brown University

World Health Organization

Genomic data, Tanzania

Genomic surveillance

- A country-wide survey was conducted in 2021 with 7666 samples from 13 regions
- R561H mutation was found in 7.4% (n=447) and 0.6% (n=302) of samples from Kagera and Tabora regions.
- In Kagera, most of the mutants were from one facility near the Rwanda border
- Prevalence of 561H increased 2021 to 2022 in Kagera but not 2022 to 2023. Instead, a different mutation, 675V is increasing in some districts



World Health Organization

Source: Deus S. Ishengoma, NIMR

In-vivo data, Tanzania

Kagera

- The finding of high 561H prevalence prompted a special TES in Kagera in 2022.
- The study confirmed artemisinin partial resistance in Tanzania with high day 3 parasitemia
- The study showed high PCR corrected efficacy for AL and ASAQ

TES Kagera, 2022

	AL		ASAQ	
	N(%)	CI 95%	N(%)	CI 95%
PCR uncorrected				
ACPR	57(64.8)	53.9-74.7	86(97.7)	92.0-99.7
Total patients PP	88		88	
PCR corrected				
ACPR	57(96.6)	88.3-99.6	86(100)	95.8-100
Total patients PP	59		86	
New infections	27		2	
Non-determinant	2			
Parasitaemia on Day 3	11(12.5%)		17 (19.3%)	

Source: Deus S. Ishengoma, NIMR

Other sites

- AL was tested in four other sites in 2022
- Failure rate of 10.1% detected in one site
- All sites had low day 3 positivity rate and only 2 samples were found with k13 mutations

TES results (AL) 2022

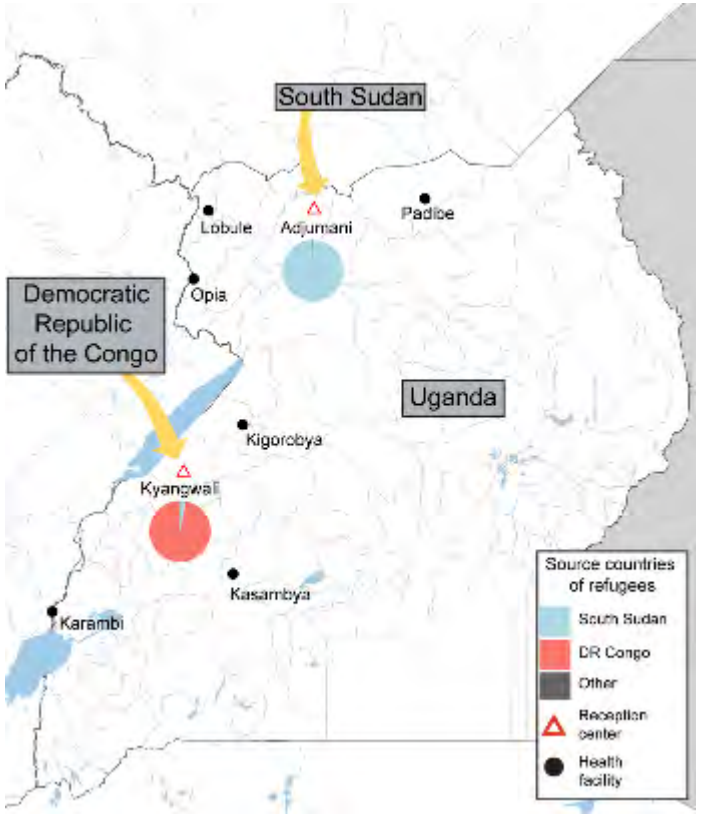
	Kigoma	Morogoro	Pwani	Tanga
Per protocol efficacy*				
% ACPR uncorrected (95% CI)	86.4 (77-92.5)	98.9 (92.9-99.9)	73.8 (62.9-82.5)	77.3 (66.9-85.2)
% ACPR corrected (95% CI)	95 (87-98.4)	98.9 (92.9-99.9)	89.9 (79.6-95.5)	94.4 (85.7-98.2)
Kaplan-Meier Efficacy				
% Uncorrected (95% CI)	86.4 (79.5-93.8)	98.9 (96.7-100.0)	73.8 (65.0-83.9)	77.3 (69.0-86.5)
% Corrected (95% CI)**	95.0 (90.3-99.9)	98.9 (96.7-100.0)	89.9 (83.1-97.3)	94.4 (89.3-99.9)

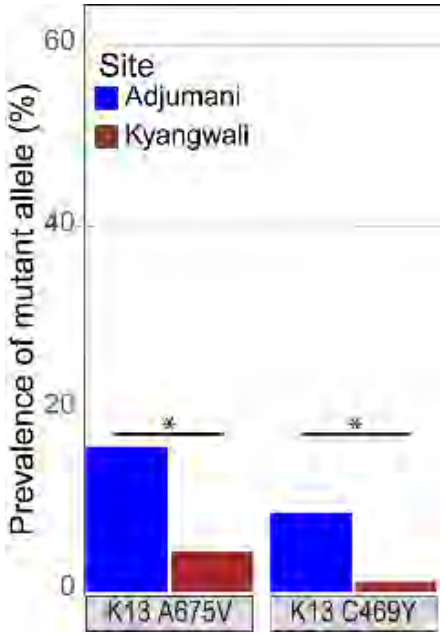
Source: Ifakara / PMI

World Health Organization

Other potential signals

- Collected samples from newly arrived refugees at settlement reception centers in Uganda show high K13 mutation prevalence among refugees in particular among South Sudanese refugees
- K13 mutations (561H & 441L) have been reported in single samples in DR Congo
- WHO is working to support a TES in South Sudan





Adjumani: refugees from South Sudan

Kyangwali: refugees from DRC

Tukwasibwe & Conrad, unpublished data

World Health Organization

Updates on activities to operationalize the strategy to respond to antimalarial drug resistance in Africa

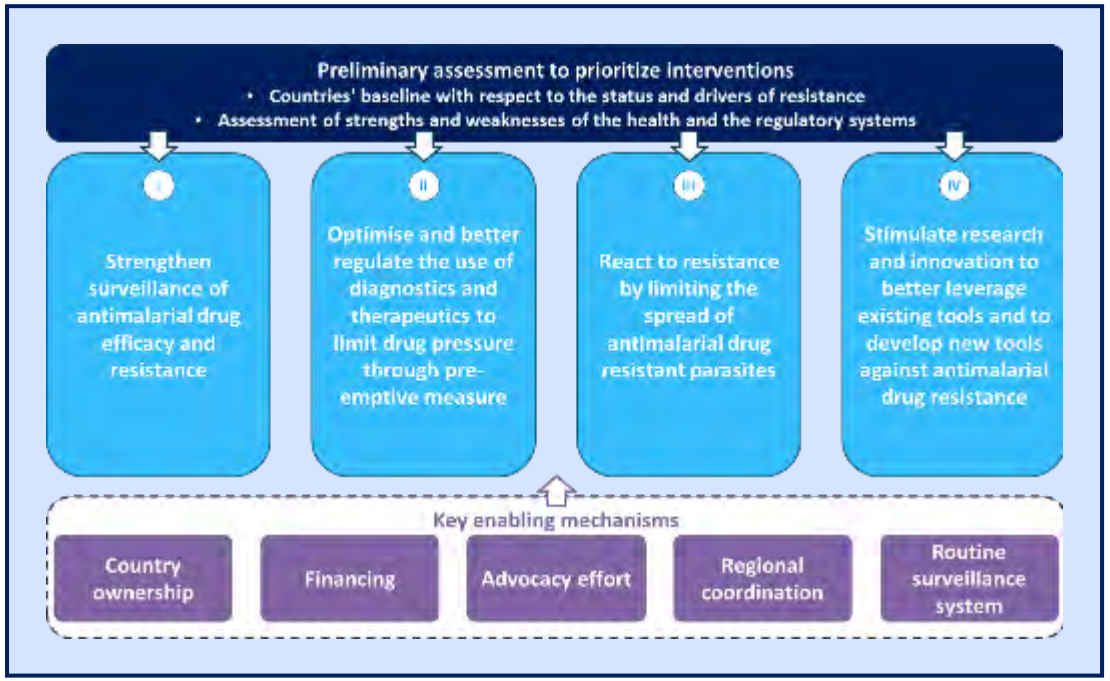
World Health Organization

- **WHO is convening two meetings** in Kampala, Uganda in November:
  - **7 & 8 November:** A regional stakeholder meeting to align on intervention priorities to support countries responding to resistance.
  - **9 & 10 November:** Meeting on surveillance of drug efficacy and resistance for countries in Eastern Africa and the Horn of Africa

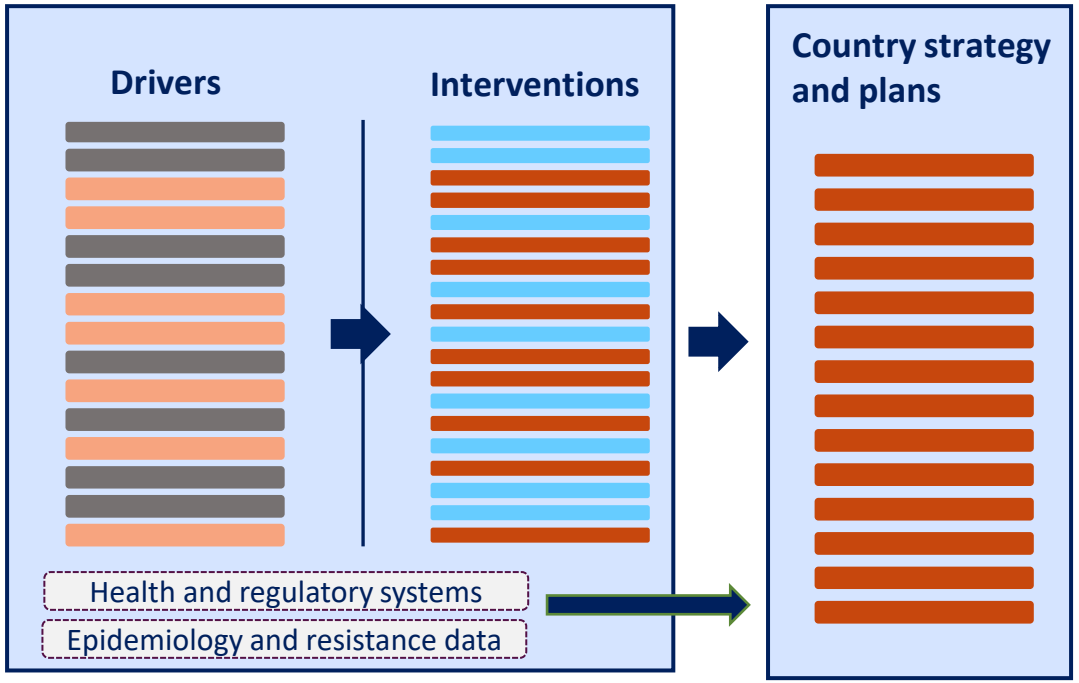
World Health Organization

Strategy to respond to antimalarial drug resistance – country assessment

- The strategy relies on better use of existing tools & development of new tools & strategies
- The starting point of the strategy was to look at potential factors that to drive the emergence and spread of resistance
- To respond to resistance, the strategy propose 20 interventions in four pillars



- To help prioritize interventions in the local context, the strategy proposes country assessments looking at factors that could drive resistance in a given context, and the systems in place and data available
- This assessment will inform country specific strategies and plans



Rwanda country assessment

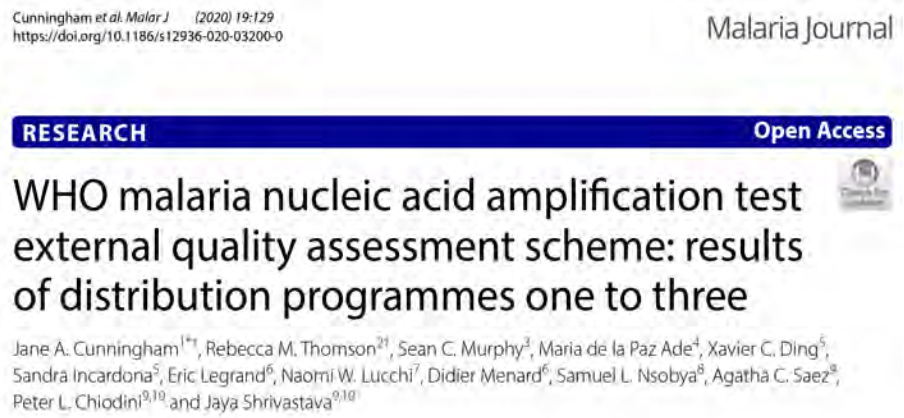
- An assessment is being done, looking at the current country situation, status of resistance and epidemiology, drivers of resistance, and overall health and regulatory systems.
- Aim is to develop assessment methodology that can be used by other countries, and present the assessment in the stakeholder meeting in November
- Once finalized, the assessment will be used to develop a country strategy including interventions to respond to antimalarial drug resistance

Emergence of antimalarial drug resistance in Rwanda
Assessment of the current country situation

World Health Organization

Expansion of the External Quality Assessment (EQA) Scheme for Molecular Markers of Antimalarial Drug Resistance

- In 2014, the MPAC endorsed the establishment of an international external quality assessment (EQA) scheme for nucleic acid amplification technique (NAAT) assays to ensure results are reliable and comparable.
- Based on the recommendations from a 2015 expert meeting, a WHO malaria NAAT EQA scheme was established in collaboration with UK National External Quality Assessment Service (NEQAS) to assess the quality of commonly used molecular diagnostic methods for detection of human malaria
- 80 labs are currently enrolled in the scheme with laboratories continually enrolling
- It allows laboratories to assess their performance by species and sample type and determine where their weaknesses may lie
- The trends of performance by submission number clearly shows an improvement in performance over time, with the weaker laboratories at the start of the scheme showing the most marked improvement in performance.



Cunningham et al. Malar J (2020) 19:129
https://doi.org/10.1186/s12936-020-03200-0

Malaria Journal

RESEARCH Open Access

WHO malaria nucleic acid amplification test external quality assessment scheme: results of distribution programmes one to three

Jane A. Cunningham[1*†], Rebecca M. Thomson[2†], Sean C. Murphy[3], Maria de la Paz Ade[4], Xavier C. Ding[5], Sandra Incardona[5], Eric Legrand[6], Naomi W. Lucchi[7], Didier Menard[6], Samuel L. Nsobya[8], Agatha C. Saez[9], Peter L. Chiodini[9,10] and Jaya Shrivastava[9,10]

	Number of labs	Number of countries
Africa	24	13
Asia	19	12
Europe	9	6
North America	11	2
South/ Central America	13	8
Australia	3	1

EQA need for molecular markers of antimalarial drug resistance

- Currently, no EQA exist for malaria drug resistance markers and molecular correction genotyping method
- As partial resistance to artemisinin is evolving in Africa there is an urgency to make sure molecular data are accurately collected and reported

➢ A virtual consultation on Expansion of the External Quality Assessment Scheme for Molecular Markers of Antimalarial Drug Resistance, July 14, 2023

The objectives of the consultation were to:

I. Agree on AM resistance markers to be prioritised for inclusion in the scheme

II. Identify EQA materials and panels needed for an expanded NAAT EQA scheme

III. Reach consensus on the functioning of the scheme (costing, implementing partners, capacity building and timeline)

World Health Organization

Conclusion from the consultation on expansion of the EQA scheme

- There was wide support for expanding the NAAT EQA scheme for inclusion of drug resistance markers and molecular correction method.
- The resistance marker EQA scheme cannot be directly incorporated into the existing EQA scheme but would be created an 'EQA plus EQA resistance marker' scheme. This would save on resources and costs.
- *PfK13* markers are the most important markers to include in the panels from the start.
- Wide support for including molecular correction methodology from early on in the expanded EQA scheme.
- Partner drug resistance markers are not considered as an immediate priority, and they may be considered for inclusion in the future when suitable validated markers become available for important partner drugs.
- The scheme has to remain flexible in terms of what resistance markers to include in the panels, as the epidemiological landscape is fluid and constantly changing and so the scheme will need to adapt to remain relevant to the situation at the time.
- An immediate step should be to conduct a survey of laboratories in the existing EQA scheme, and then pilot the new scheme among members at the meeting

Strategy implementation | Selected WHO planned and ongoing activities

- Collection and sharing of data in the **Malaria Threat Maps**

Malaria Threat Maps
(https://www.who.int/teams/global-malaria-programme/surveillance/malaria-threats-map)

Finished and published data

Planned and ongoing studies

World Health Organization

Thank you

For more information, please contact:
Charlotte Rasmussen
Diagnosis, Medicine and Resistance Unit, Global Malaria Programme
rasmussenc@who.int

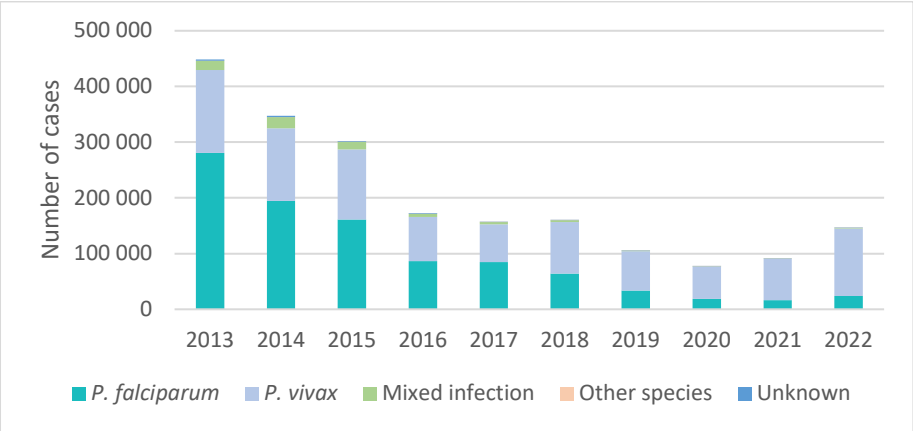
World Health Organization

The Mekong Malaria Elimination Programme

In 2013, the six Greater Mekong subregion (GMS) countries of Cambodia, China (Yunnan province), Lao People's Democratic Republic, Myanmar, Thailand and Viet Nam were faced with a momentous challenge. Multidrug resistance threatened to make *Plasmodium falciparum* malaria untreatable in the GMS, and potentially globally if resistant parasites were to escape the region and spread. The World Health Organization (WHO) developed an initial emergency response in 2013–2015 to contain resistance and address the *P. falciparum* malaria burden. However, to address the continuous threat posed by antimalarial drug resistance, a region free of malaria was envisaged in the WHO *Strategy for malaria elimination in the Greater Mekong subregion: 2015–2030 (1)*. This strategy was supported by the Ministerial Call for Action to Eliminate Malaria in the GMS before 2030 *(2)*, signed by the Ministers of Health of all the GMS countries in 2018. Since this call to action, sustained political momentum and extensive community mobilization have supported dramatic reductions in malaria case numbers and deaths across the region.

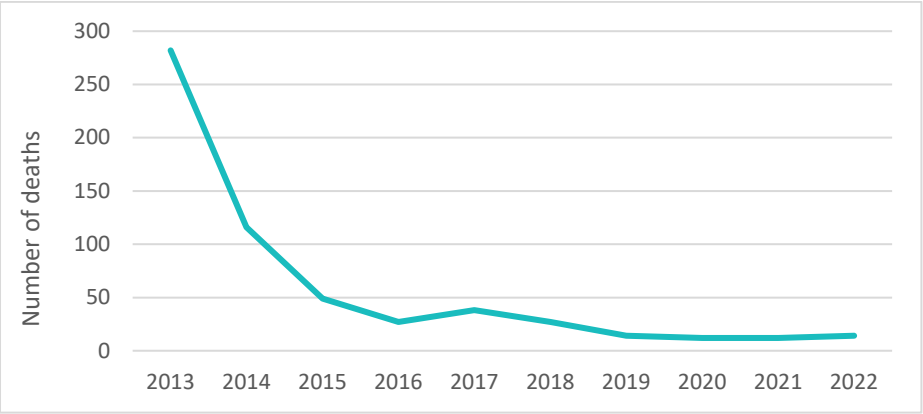
Over the last nine years, the GMS countries have made remarkable progress towards their collective goals of *P. falciparum* elimination by 2023 and elimination of all human malaria species by 2030. In 2013, there were 448 247 confirmed malaria cases in the GMS, compared to 146 718 in 2022 – a 67% reduction in cases overall. *P. falciparum* and mixed cases have declined from 297 998 to 25 105 over the same period – a 92% reduction – and deaths due to malaria have decreased by 95% (Figs. 1 and 2) *(3)*.

Fig. 1. Malaria cases in the GMS, 2013–2022



Source: Mekong Elimination Database *(3)*

Fig. 2. Malaria deaths in the GMS, 2013–2022



Source: Mekong Elimination Database *(3)*

The coronavirus disease (COVID-19) pandemic presented a major threat to malaria elimination efforts, disrupting malaria services and elimination activities, even though restrictions on population movement reduced malaria transmission. Once social restrictions were eased, rapid reinstatement of capacity for malaria control, diagnosis and treatment prevented the anticipated rebound in cases, with progress towards malaria elimination reverting to pre-pandemic trends in 2022.

The Greater Mekong subregion (GMS) countries are strongly committed to malaria elimination and cases continued to decline in most areas in 2022. However, the continued unstable political situation in Myanmar caused an overall increase in cases across the GMS to 142 777 in 2022, compared to 91 048 in 2021. Although most of this increased burden occurred in Myanmar, the border regions of neighbouring countries, in particular Thailand, were also affected.

Reaching the unreached populations, particularly in remote and marginalized communities, is vital for malaria elimination. This requires a comprehensive, tailored and participatory approach that considers social, economic and political factors, with collaboration among stakeholders to ensure access to prevention and treatment interventions. The role of community-based volunteer health workers is especially important in gaining trust and understanding needs. Civil society organizations are also an important component in fostering community engagement and ownership.

Despite the presence of partial artemisinin resistance in the GMS, several artemisinin-based combination therapies remain highly effective against *Plasmodium falciparum*. The number of *P. falciparum* and mixed cases increased from 17 115 in 2021 to 25 105 in 2022, although the proportion of all cases that were caused by *P. falciparum* declined from 18.8% to 16.8% over the same period.

P. vivax is the dominant parasite in the region, causing 83% of cases in 2022 and presenting a significant barrier to malaria elimination. The number of *P. vivax* cases increased from 73 856 in 2021 to 121 309 in 2022, and effective strategies for addressing *P. vivax* malaria elimination are urgently needed.

As elimination goals are approached, high-quality epidemiological data are needed to identify and address transmission foci, particularly across country border zones. The Malaria Elimination Database continues to foster collaboration and facilitate data sharing and epidemiological monitoring, supporting strategic decision-making, coordination and communication across the GMS.

Integrated drug efficacy surveillance is being implemented in areas where malaria case incidence has sufficiently declined in order to enable comprehensive follow-up of every malaria case and ensure treatment completion, while monitoring antimalarial effectiveness.

Countries in the GMS are preparing for national malaria-free certification, with subnational verification serving as a valuable programmatic exercise to support compliance with processes and

documentation. Planning for prevention of re-establishment of malaria is essential to fulfil the criteria for malaria-free status.

Given anticipated reductions in donor funding for malaria, ensuring the sustainability of malaria elimination programmes in the GMS is essential. GMS countries are actively developing transition plans to shift towards domestically financed and supported malaria responses.

References

1. Strategy for malaria elimination in the Greater Mekong Subregion: 2015–2030. Manila: World Health Organization Regional Office for the Western Pacific; 2015 (https://apps.who.int/iris/handle/10665/208203, accessed 8 August 2023).
2. Amid concern over drug resistance, Mekong countries call for accelerated action to eliminate malaria before 2030 [website]. Geneva: World Health Organization; 2017 (https://www.who.int/southeastasia/news/detail/08-12-2017-amid-concern-over-drug-resistance-mekong-countries-call-for-accelerated-action-to-eliminate-malaria-before-2030, accessed 8 August 2023).
3. Mekong Elimination Database [website]. Phnom Penh: World Health Organization (http://medb-gms.org/, accessed 8 August 2023).

MALARIA FREE MEKONG

Progress of malaria elimination in the GMS

Dr Pascal Ringwald, coordinator Mekong Malaria Programme

MEKONG MALARIA ELIMINATION PROGRAMME

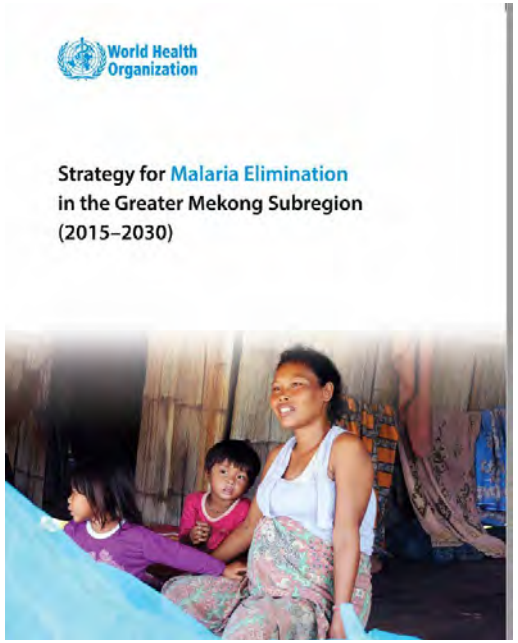
World Health Organization

Rational for malaria elimination in the GMS

- Artemisinin resistance has emerged independently in multiple geographic areas within the GMS, raising concerns about effectiveness of a “firewall approach”;
- Multidrug resistance including ACT resistance was reported in the GMS;
- The burden of disease in the GMS has been lowered to levels where most countries are considering, or have already committed to, elimination over the next 10–15 years;
- *P. falciparum* elimination in the GMS appears technically and operationally feasible at a reasonable cost

Success of malaria elimination in the GMS

- GMS countries agreed on a common strategy
- Political support (ministerial call of action signed at WHA 2018)
- Financial support (mainly through GF RAIE)
- Oversight committee (Regional Steering Committee)
- Technical support and coordinatization of partners lead by WHO



MEKONG MALARIA ELIMINATION PROGRAMME

World Health Organization

Malaria elimination targets in the GMS

China

30 June 2021, China certified as malaria-free

2021

***P. falciparum* target of GMS**

P. falciparum malaria eliminated in all countries of the GMS (except Myanmar 2026)

2023

Thailand Target

All malaria species eliminated in Thailand

2024

Cambodia Target

All malaria species eliminated in Cambodia

2025

GMS Target

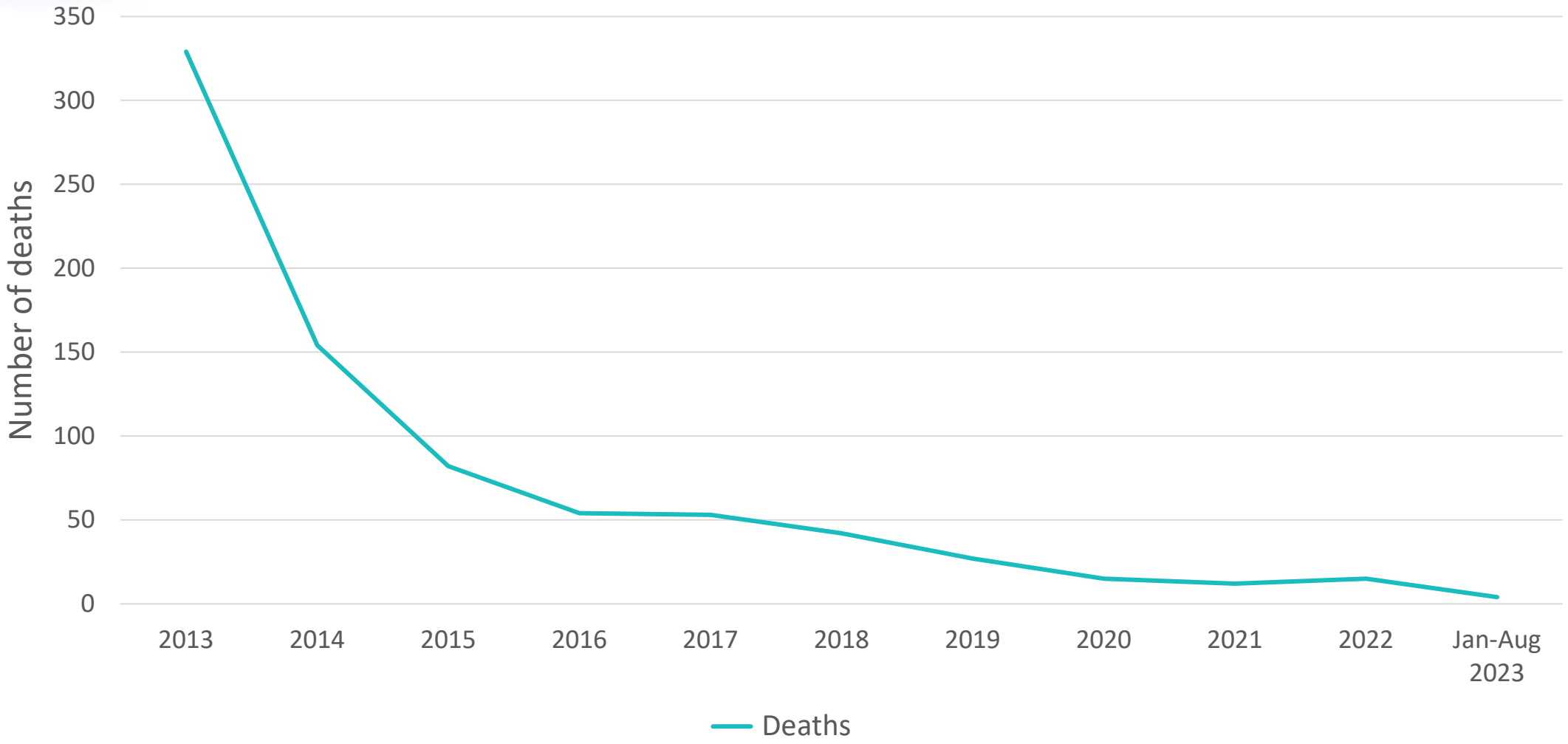
All malaria species eliminated in all countries of the GMS

2030

MEKONG MALARIA ELIMINATION PROGRAMME

World Health Organization

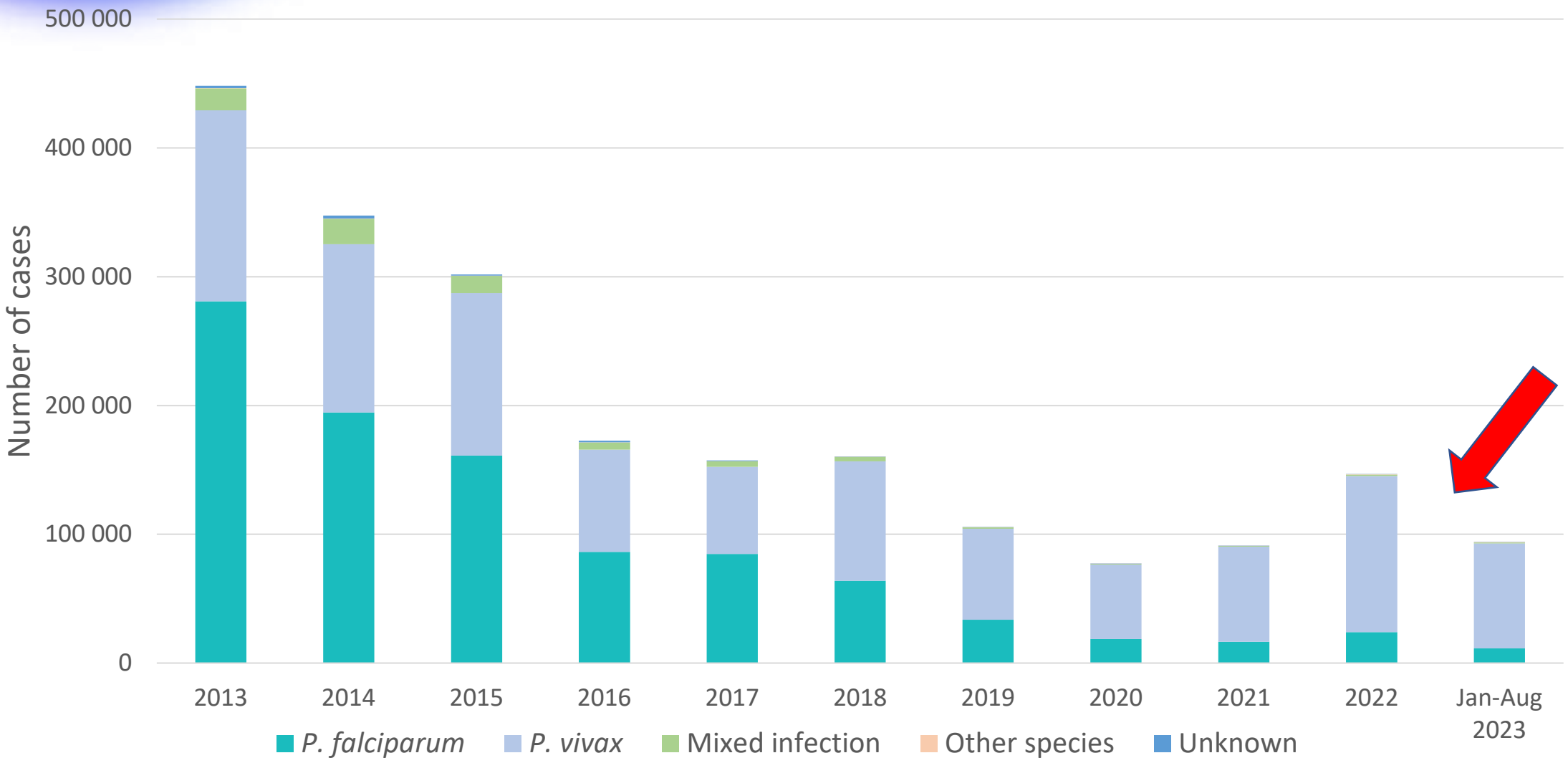
Reduction of malaria death in the GMS



MEKONG MALARIA ELIMINATION PROGRAMME

World Health Organization

Progress in malaria elimination in the GSM

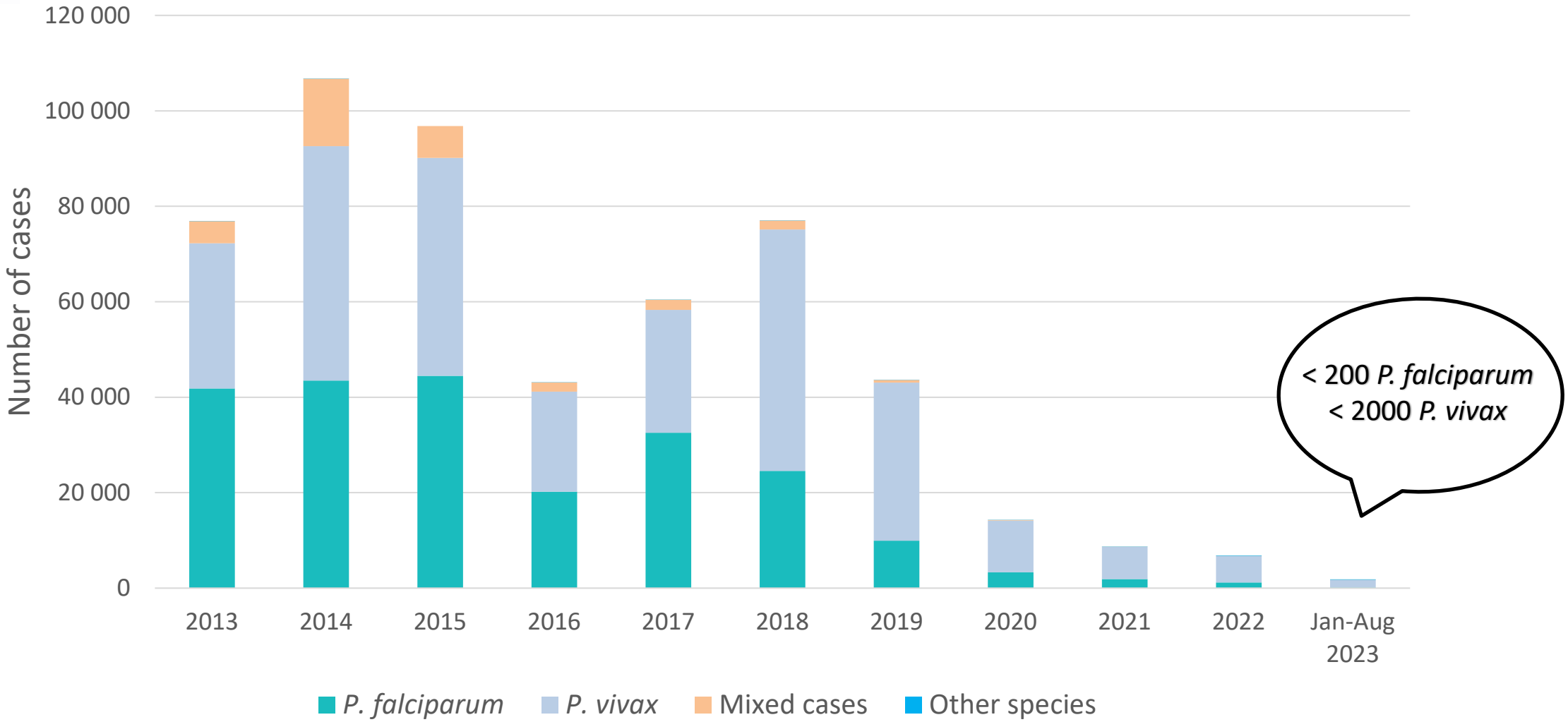


MEKONG MALARIA ELIMINATION PROGRAMME

Note: Myanmar of 2022 is only partial data
Data source: MEDB database

World Health Organization

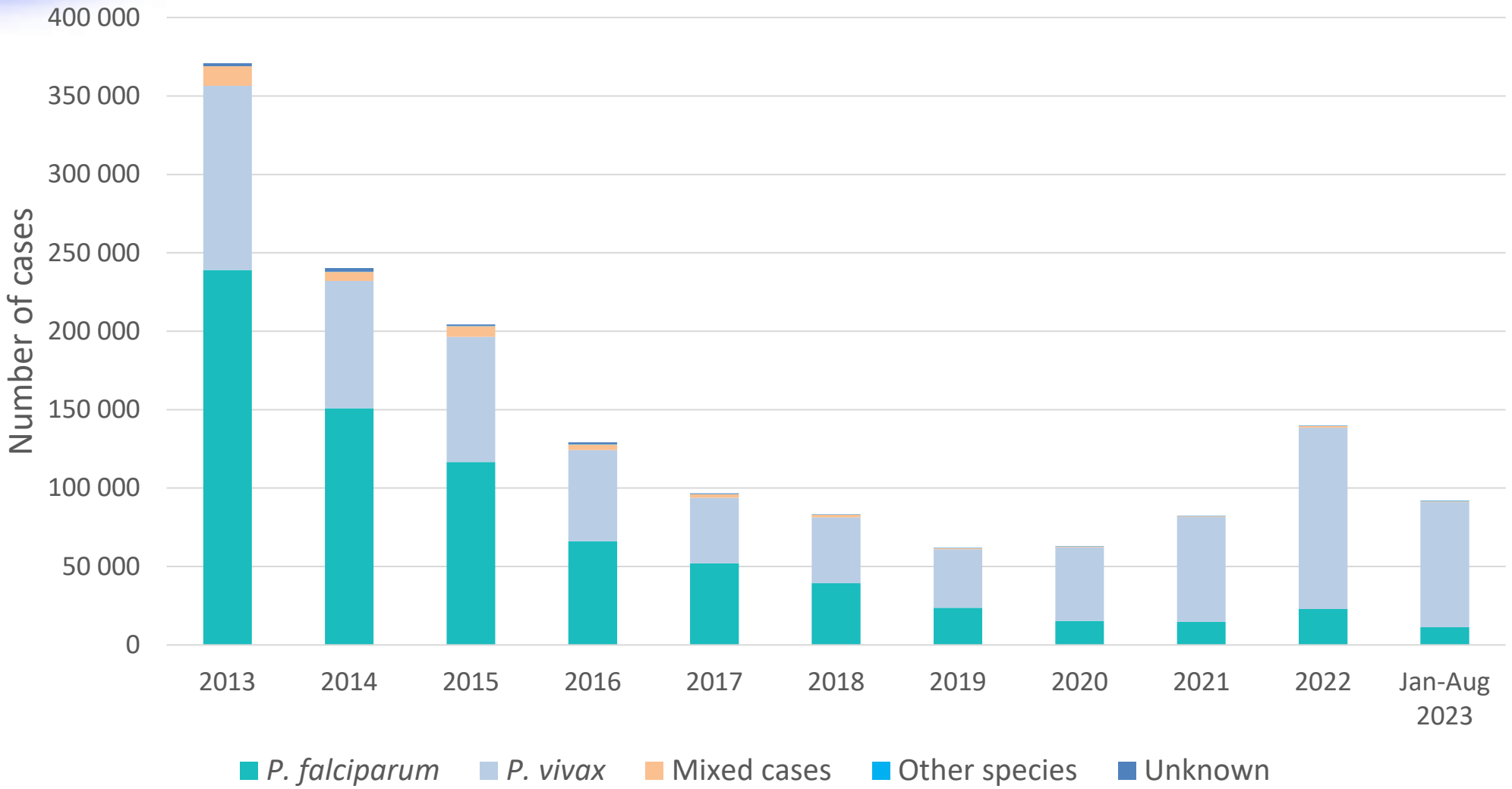
Progress in malaria elimination in WPRO



MEKONG MALARIA ELIMINATION PROGRAMME

World Health Organization

Progress in malaria elimination in SEARO

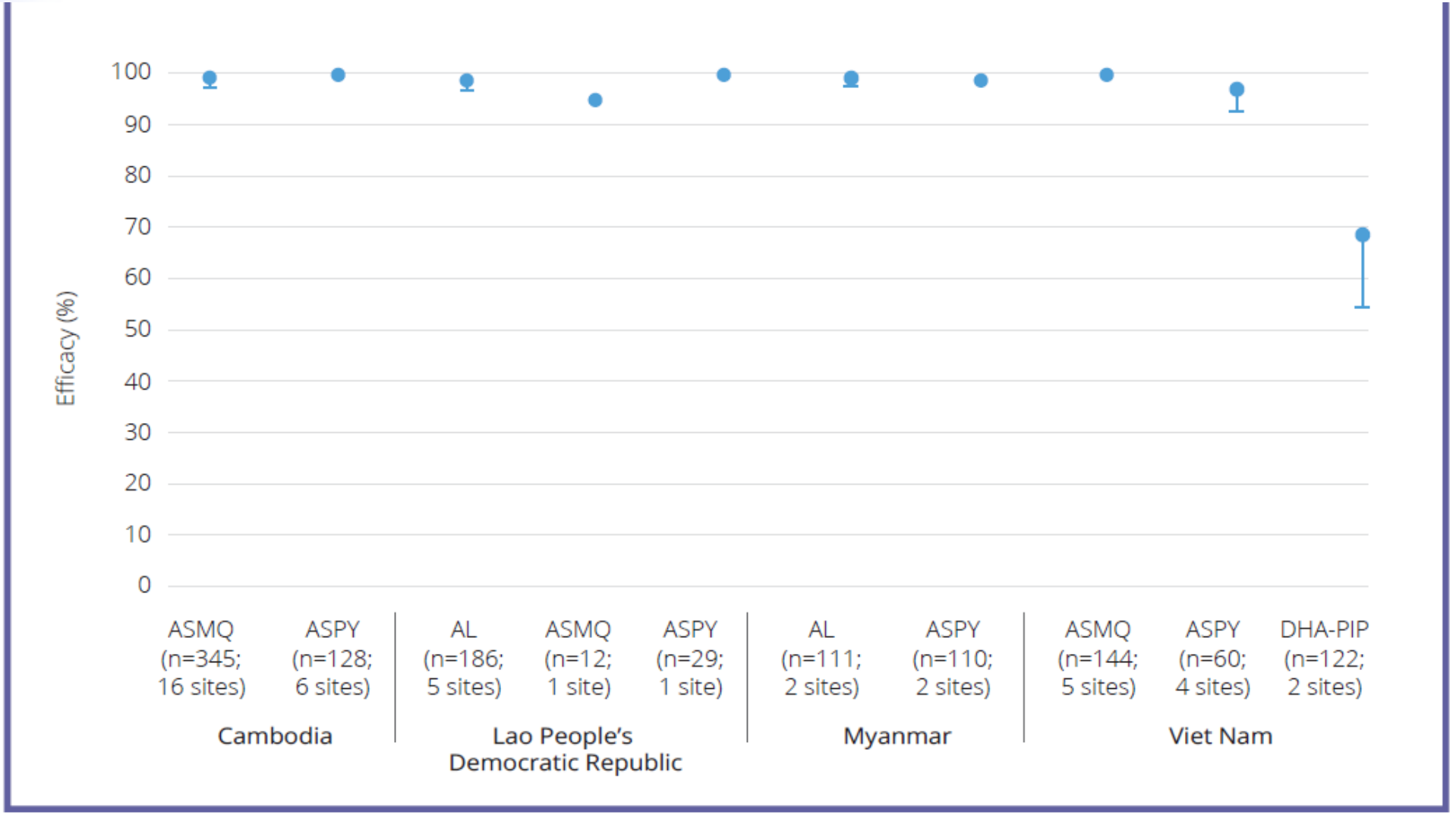


MEKONG MALARIA ELIMINATION PROGRAMME

Note: Myanmar of 2022 is only partial data
Data source: MEDB database

World Health Organization

Drug efficacy in the GMS between 2018 and 2022



MEKONG MALARIA ELIMINATION PROGRAMME

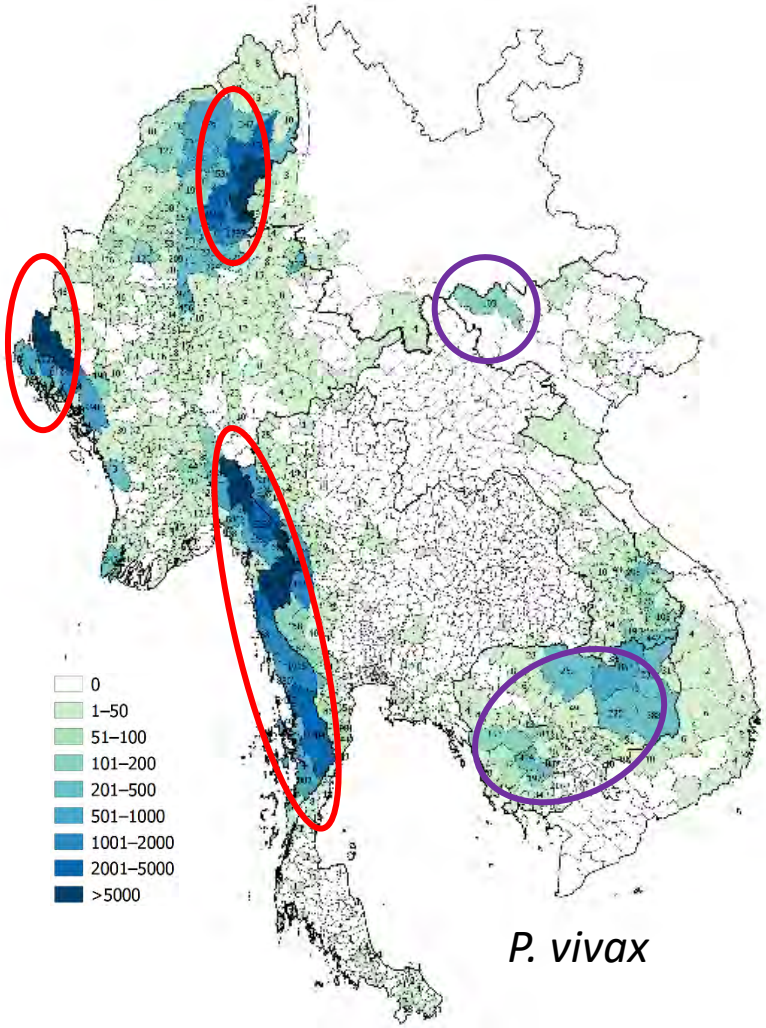
World Health Organization

Number of ACTs with efficacy > 90% in the GMS

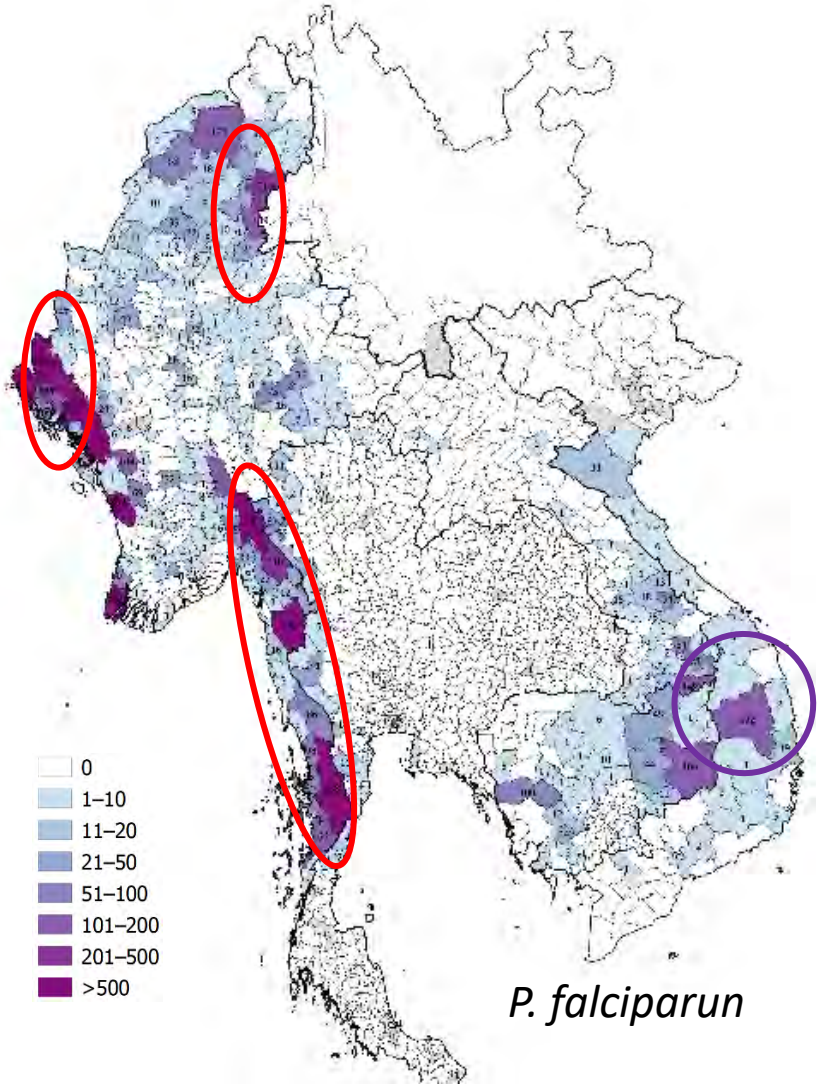
MEKONG
MALARIA
ELIMINATION
PROGRAMME

World Health Organization

Distribution of malaria in the GMS



P. vivax

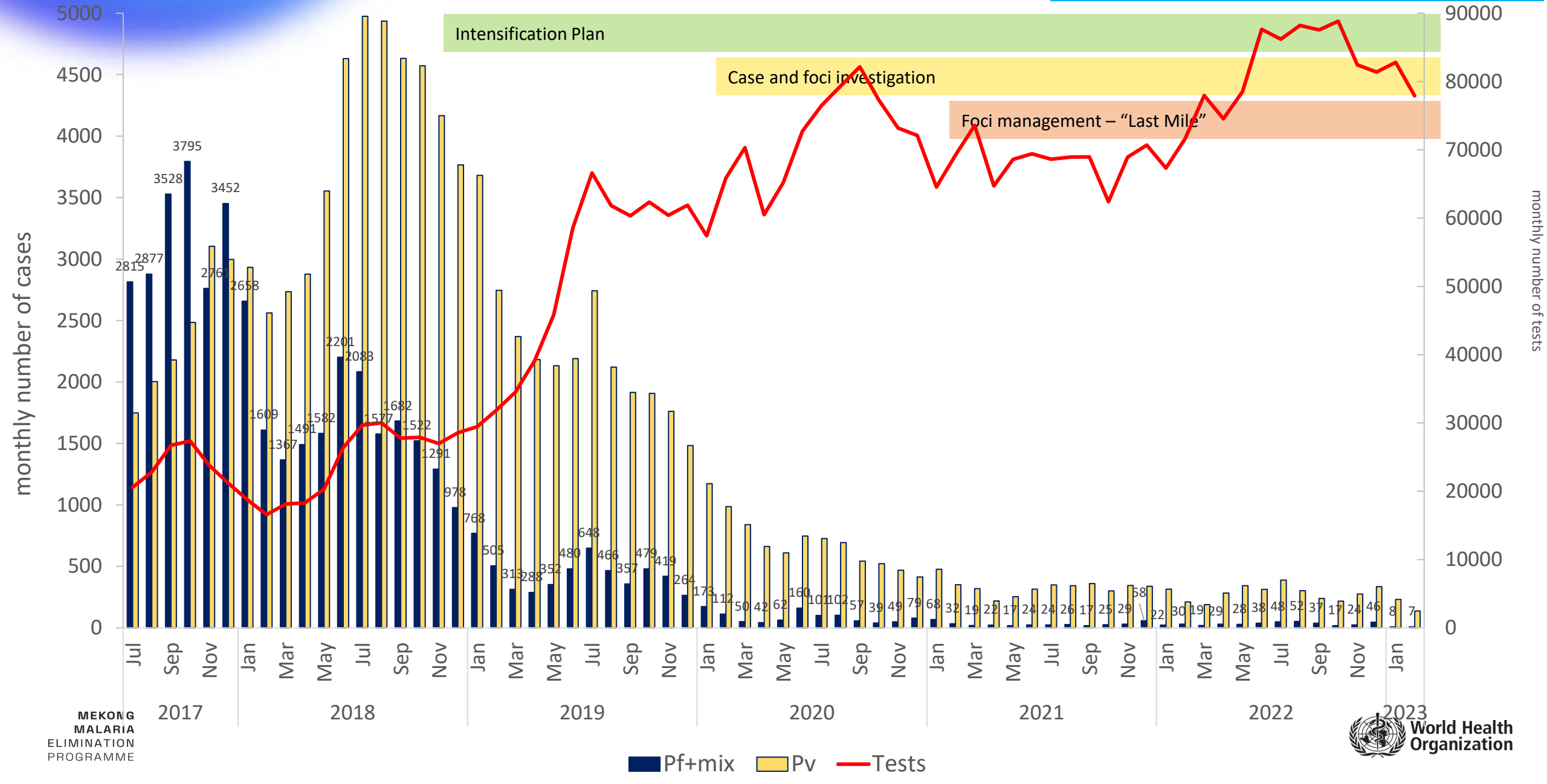


P. falciparun

MEKONG
MALARIA
ELIMINATION
PROGRAMME

World Health Organization

Example of Cambodia



MEKONG MALARIA ELIMINATION PROGRAMME

World Health Organization

Intensification plan

- **Decentralized technical support** with epidemiologists in targeted provinces

- Intensify malaria activities based on the monthly epidemiological analyses and mapping of hot-spots:
  - **Village malaria workers (VMWs)** deployed in hotspot areas;
  - **Mobile malaria workers (MMWs)** deployed in high-risk forested areas;
  - **Active case detection** in the forests to ensure early detection and treatment;
  - **Distribution** of forest packs to forest goers including **hammock, LLIHN, boots and backpack.**

Intensification plan

Case and foci Investigation

- Cambodia rolled out investigations in 2020 to detect, investigate and clear all cases and foci
- **Case investigations** were led by health centers for all *P. falciparum* cases: classified based on travel
- In 2021, the case investigation form simplified to enable **village malaria workers (VMWs)** to conduct case investigations
- **Mobile malaria workers (MMWs)** operating in high-burden areas

MALARIA CASE INVESTIGATION FORM ONLY FOR *P. Falciparum or mixed*

SECTION 1: CASE NOTIFICATION BASED ON REPORT FROM PLACE OF DIAGNOSIS

Date of notification: / / Case ID#
Family name: Given name:
Telephone: Age (Years):
Sex: M ☐ F ☐
Village of residence: HC: OD: Province:

Type of Notification
Passive case detection ☐ Re-active case detection ☐ Pro-active case detection ☐
Point of Care OD Point of Care Province

If Passive case detection
Referral Hospital ☐
Former District Hospital ☐
Health Center ☐
Health Post ☐
Village Malaria Worker ☐
Mobile Malaria Worker ☐
Private provider ☐
Armed Force ☐
Police ☐
Point of care name
Point of care ID #

If Re-active case detection
Index case ID # *System Generated*
Health Center name
Point of care ID #

If Pro-active case detection
Mobile and migrants screening ☐
Border screening ☐
Focal mass screening ☐
Focal fever screening ☐
Focal targeted screening ☐

SECTION 2: CASE INVESTIGATION
Date of investigation: / /
Conducted by (name): Telephone #:
Introduction to village authority, community leaders Village with VMW ☐ Available today for case investigation ☐
Localization of the residence: GPS coordinates LONG LAT
Verify case ID, introduction, informed verbal consent Respondent: Case himself ☐ Relative ☐ Not met ☐

CASE'S PROFILE
Occupation: Agriculture, farming ☐ Manufacture ☐ Student ☐ Trade, service ☐ Civil servant ☐ Other ☐
Current residence since : > 1 year ☐ 6 months - 1 year ☐ 1 week - < 6 months ☐ < 1 week ☐
Citizenship: Cambodia ☐ Other country:

HISTORY OF CURRENT EPISODE
Symptoms before diagnosis: Fever ☐ Chills ☐ Sweat ☐ Headache ☐ Nausea ☐ Vomiting ☐ Diarrhoea ☐
Date of first symptoms / / No symptom ☐ Other notifiable signs

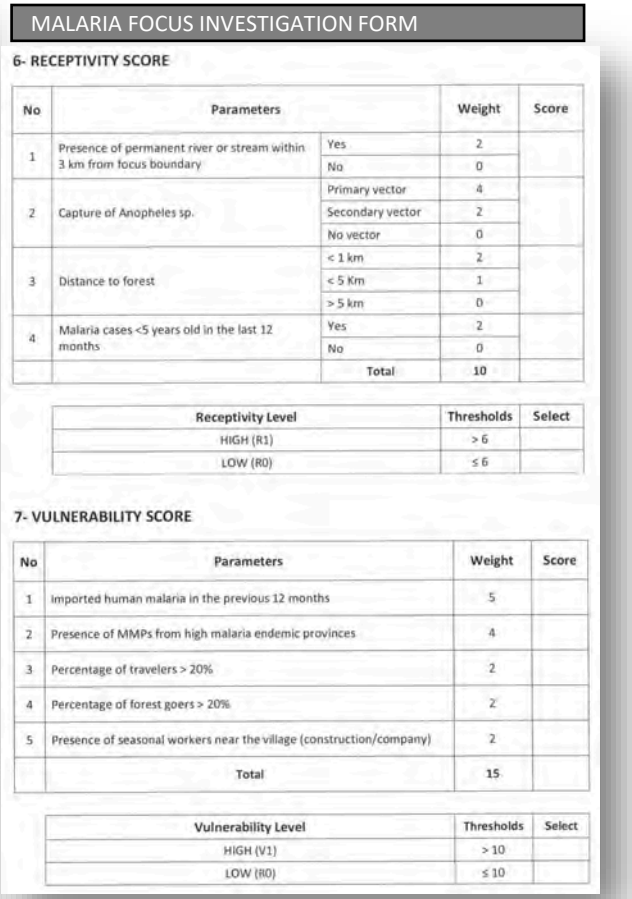
COMPLETION OF TREATMENT **to check used blister/ box if available**
Treatment prescribed: AS+MQ ☐ Other SLD Primaquine ☐ ____ mg
Treatment completed ☐: Tablets of ____ mg / ____ mg ____ tablets ____ times per day, for ____ days
Treatment not yet completed ☐ (if visited before D2)

MALARIA HISTORY
Had malaria ever: Yes ☐ No ☐ If yes, date last episode of malaria: / /
Was confirmed by testing ☐ Diagnosis made by: Public HF ☐ VMW ☐ PPM ☐ Other ☐
Got treatment from: Public HF ☐ PPM ☐ VMW ☐ Pharmacy ☐ Shop ☐
Remember which drug? for ____ days?

HOUSE-HOLD AND PREVENTION (Optional)
Did somebody in your household had malaria ☐ Ever Within last month ☐ Last 12 months ☐
How many people living in this house-hold? ____ How many separate sleeping places? ____
How many mosquito nets ____ Did you sleep under a mosquito net last night? Yes ☐ No ☐
Got the net less than one year ago ☐ 1-2 year ☐ More than 2 years ☐ More than 3 years ☐
Got the net from Government ☐ From NGO ☐ Bought it from shop/market ☐ (HE on spot)

SECTION 3: CASE CLASSIFICATION
1. Did you sleep **every night** in this village within the last 2 weeks?
   Yes ☐ → **classify L1 CONDUCT REACTIVE CASE DETECTION**
   No ☐ → Go to 2
2. Did you sleep **at least one night** in another village in the **same catchment area of this HC** within the last 2 weeks?
   Yes ☐ → **classify L2 AND CONDUCT REACTIVE CASE DETECTION** Name of the other village
   No ☐ →Go to 3
3. Did you sleep **at least one night** in another village in the **same OD** within the last 2 weeks?
   Yes ☐ → **classify L3** No ☐ →Go to 4
4. Did you sleep **at least one night elsewhere in Cambodia** within the last 2 weeks?
   Yes ☐ → **classify L4** In which province? No ☐ → Go to 5
5. Did you sleep **at least one night in another country** within the last 2 weeks?
   Yes ☐ → **classify IMPORTED** In which country? No ☐ → Go back to 1

FOR L1/L2 CASES ONLY: DETAIL OVERNIGHT STAYS AROUND THE VILLAGE
Did you sleep outside your house around the village? Yes, last week ☐ Yes, the week before ☐ No ☐ → Stop here
Did you sleep in the forest? No ☐ Yes ☐ → Harvesting ☐ Logging ☐ Hunting ☐ Fishing ☐ Other ☐
Did you sleep at worksite? No ☐ Yes ☐ → Plantation ☐ Farm ☐ Logging ☐ Mine ☐ Construction site ☐
Did you sleep in: A house? ☐ A plot hut? ☐ A tent? ☐ A camp? ☐
Did you sleep under: A mosquito net? Yes ☐ No ☐ A hammock with net? Yes ☐ No ☐



MALARIA FOCUS INVESTIGATION FORM

6- RECEPTIVITY SCORE

No	Parameters		Weight	Score
1	Presence of permanent river or stream within 3 km from focus boundary	Yes	2	
		No	0	
2	Capture of Anopheles sp.	Primary vector	4	
		Secondary vector	2	
		No vector	0	
3	Distance to forest	< 1 km	2	
		< 5 Km	1	
		> 5 km	0	
4	Malaria cases <5 years old in the last 12 months	Yes	2	
		No	0	
		Total	10	

Receptivity Level	Thresholds	Select
HIGH (R1)	> 6	
LOW (R0)	≤ 6	

7- VULNERABILITY SCORE

No	Parameters	Weight	Score
1	Imported human malaria in the previous 12 months	5	
2	Presence of MMPs from high malaria endemic provinces	4	
3	Percentage of travelers > 20%	2	
4	Percentage of forest goers > 20%	2	
5	Presence of seasonal workers near the village (construction/company)	2	
	Total	15	

Vulnerability Level	Thresholds	Select
HIGH (V1)	> 10	
LOW (R0)	≤ 10	

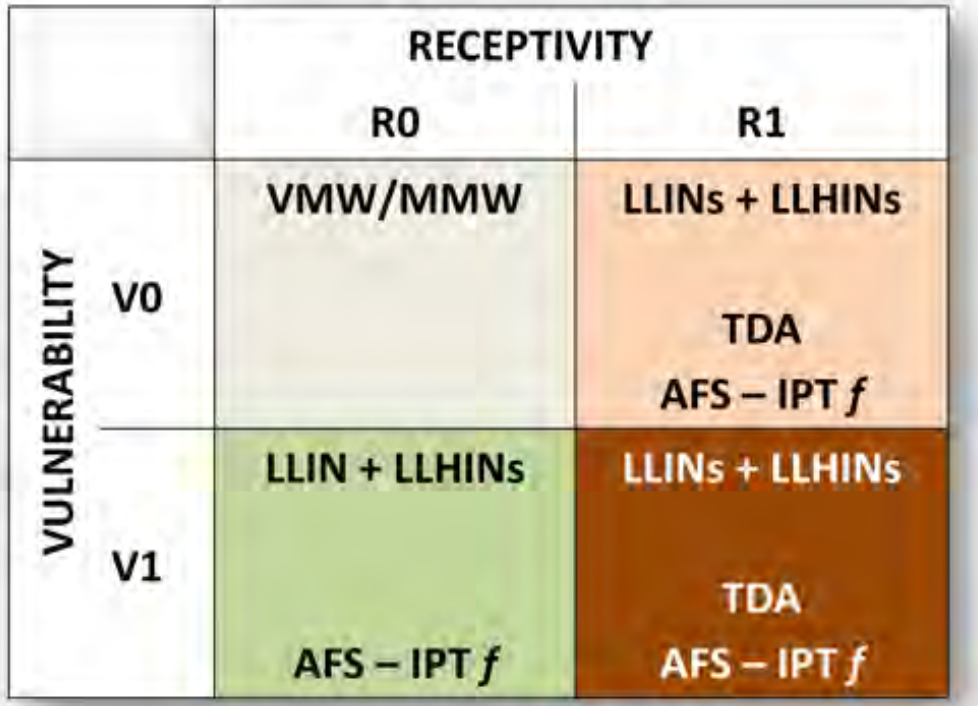
Pictured: case and foci investigation forms from 2020 (forms have changed slightly since, to allow VMWs to conduct case investigations)

MEKONG MALARIA ELIMINATION PROGRAMME

World Health Organization

Foci management: intensification approaches

- Based on receptivity and vulnerability scores from foci classification and foci management:
  - Vector control: LLINs & LLIHNs
  - Rigorous case management: weekly house-to-house fever-screening, active fever screening
  - More focalized approaches to accelerate malaria elimination:
    - Targeted drug administration (TDA) for adult males aged 15-49
    - Intermittent preventive treatment (IPTf) for travelers to high-risk areas (forest)



Challenges in malaria elimination

- **Disruptions to the supply chain** of antimalarial drugs and diagnostic tools can impede effective malaria control.
- **Procurement delays and wastage**: As malaria cases decrease and elimination progresses, there may be challenges in procuring enough antimalarial medicines.
- **Integration of VMWs** into primary health care and universal health coverage systems is vital for sustainable malaria control.
- **The situation in Myanmar** adds another layer of complexity, and tailored interventions and increased vigilance are needed to both regain control of the malaria situation in Myanmar and minimize the impact on neighbouring countries.
- ***P. vivax*** causes most of the malaria cases in the GMS, and implementing and scaling up radical cure coverage is essential to reduce the malaria burden and drain the transmission reservoir.
- **Ensuring access to G6PD testing** remains a challenge, especially for hard-to-reach populations.

Challenges in malaria elimination (2)

- **Zoonotic malaria caused by *P. knowlesi*** requires specific surveillance and control strategies.
- **Reaching hard-to-reach populations**: Achieving universal coverage of malaria diagnosis and treatment services is essential for malaria elimination.
- **Loss of political interest** and insufficient political engagement: Sustaining the momentum of malaria elimination efforts requires unwavering commitment and support from the highest levels of government.
- **The need for cross-border collaboration**: Malaria knows no boundaries, and effective cross-border collaboration is vital in addressing outbreaks and epidemics.
- **Sustainability of malaria surveillance**, control, diagnosis and treatment: Sustainable funding is essential to support ongoing malaria elimination efforts.

MEKONG
MALARIA
ELIMINATION
PROGRAMME

MALARIA FREE MEKONG

MEKONG
MALARIA
ELIMINATION
PROGRAMME

World Health Organization

24th meeting of the Malaria Policy Advisory Group

31 October 2023

Dr Daniel Ngamije, Director, WHO Global Malaria Programme

World Health Organization

New GMP operational strategy, 2024-2030

- Strategy recognizes unique role of WHO/GMP. It identifies how GMP can do its core business better and also be more transformative in addressing emerging and pressing issues, subject to sufficient resources.
- Success will depend on working with other malaria partners to address key implementation challenges and achieve country-level impact.
- GMP welcomes comments from partners and looks forward to working with all stakeholders to ensure a more impactful response
- Strategy will enhance – and not replace – a country's national strategic plan
- GMP is actively interacting with developers, e.g. by sharing PPCs and TPPs to help developers understand what the needs are
- Effective implementation of the strategy will be underpinned by:
    - Strong health systems
    - Community ownership
    - A multi-sectoral approach

World Health Organization

Guiding principles for prioritizing malaria interventions

- Strong appetite for this guidance among malaria stakeholders
- Important to align “guiding principles” with “subnational tailoring” documents
- Guidance must be responsive to the needs of NMCPs, and additional NMCP input was recommended
- Acknowledgement of data gaps / insufficient capacity to undertake surveillance in many countries
- Some stakeholders pointed out a tension between the need for guidance and being “too prescriptive”
- Not just an exercise to cut costs, but also an opportunity to make the case for additional funding

World Health Organization

Comparative effectiveness in the context of the arrival of new vector control products

- Significant progress made within GMP to address MPAG recommendations from the last meeting.
  - Comparative efficacy evaluation mainstreamed into guidelines development process
  - Further progress made to coordinate GMP and PQ processes, with the overall aim of reducing time to market for new vector control products.
- Broader WHO initiative aligning processes for prequalification and guidelines development. GMP is playing a role in this process and helping inform best practices.

World Health Organization

High burden to high impact (HBHI) approach

- The evaluation demonstrated positive impact in 4 countries (Cameroon, Ghana, Niger, Mali) with regards to:
    - o improved use of strategic information and data for decision-making
    - o high-level political will / elevation of malaria programme within MoH
    - o shift from “one-size-fits-all” to tailored approaches that optimize impact
- The evaluation demonstrated challenges related to:
    - o lack of awareness at the local and district levels
    - o generally perceived as a “project” and not an approach
    - o diversion of resources due to the COVID-19 pandemic
    - o no guidelines for implementation
    - o no M&E framework to track implementation
    - o no resources for socialization of the HBHI approach

World Health Organization

Response to antimalarial drug resistance in Africa

- Antimalarial drug resistance is a major threat to future progress against malaria in Africa – we need to act with urgency and determination
- WHO will continue to work with countries to translate the *Strategy to respond to antimalarial drug resistance in Africa* into tangible actions.
- **Short term actions:** introduction of multiple first-line drug therapies, including market shaping; strengthening surveillance of drug efficacy and resistance; and addressing the over-use of monotherapies.
- **Longer-term actions:** investment in new products, accelerating new drug approvals; developing robust drug resistance surveillance networks; and enabling countries to develop resistance response capacity.
- Meeting with NMCPs on 7-10 Nov 2023 will focus on operationalizing the strategy.

World Health Organization