Statistical analysis plan for the Malaria Vaccine Pilot Evaluation (MVPE)

Version 1.0, 3 Dec 2018 (outline of analysis approach)
Version 2.1, 17 September 2019 (SAP for circulation)
Version 3.0, 11 June 2020 (comments from PAG, DSMB and investigators)
Version 3.2, 16 June 2020 (updated analysis methods)
Version 3.3, 6 Sept 2020 (updated analysis timings)
Version 3.4, 29 Jun 2021 (updated methods for denominators, contamination and confounding, comments from investigators)
Version 3.42 28 Jul 2021 (final version for month 24 analysis)

Trial registration: clinicaltrials.gov NCT03806465

Protocol: RTS,S_MVIP_v8 18jul18. An evaluation of the cluster-randomised pilot implementation of RTS,S/AS01 through routine health systems in moderate to high malaria transmission settings in sub-Saharan Africa. 18 July 2018
RTS,S MVIP v9.0. 22 Apr 2020. An evaluation of the cluster-randomised pilot implementation of RTS,S/AS01 through routine health systems in moderate to high malaria transmission settings in sub-Saharan Africa: a post-authorization observational study.

Study principal investigators and evaluation partners:
WHO, Geneva: Mary Hamel, David Schellenberg
Kintampo Health Research Centre, Kintampo, Ghana: Kwaku Poku Asante
University of Malawi, College of Medicine, Blantyre, Malawi: Don Mathanga
CDC Kenya Malaria Programme, Kisumu, Kenya: Aaron Samuels

PATH (Economic and qualitative evaluation): Scott Gordon

Analysis plan prepared by: Paul Milligan, LSHTM

Contents

1. Background

The Malaria Vaccine Pilot Evaluation (MVPE) [1] is a joint initiative of the World Health Organisation's Department of Immunization, Vaccines and Biologicals, and the Global Malaria Programme, to evaluate the effectiveness of the RTS,S/AS01 malaria vaccine delivered through the routine immunisation services in Kenya, Malawi and Ghana, over 4 years.

In the phase 3 trial of the vaccine conducted from 2009 to 2014 [2], in children who received four vaccine doses (three primary doses and a fourth dose 18 months later), the incidence of uncomplicated malaria and of severe malaria were reduced by 39% and 32% respectively over 4 years of follow-up. For every 1000 children vaccinated with 4 doses, a total of 1774 cases of malaria and 40 hospital admissions due to malaria were averted over about 4 years[1]. A total of 19 cases of severe malaria were averted, cases which, without the enhanced access to care provided in the context of the trial, might have led to death. The vaccine therefore could have a substantial public health impact.

The current vaccination schedules include three primary doses to be administered before 9 months of age and a fourth dose at about 2 years of age. This involves three to four additional vaccination visits than are currently recommended in this age range. The MVPE seeks to determine the impact realised in practice when the vaccine is introduced in areas of high to moderate malaria incidence in three countries with year-round malaria transmission. The programme will evaluate the safety of RTS,S/AS01 in routine use, the feasibility of administering four doses of the vaccine, and the impact of the vaccine at population level, using a cluster-randomized design. The vaccine is being introduced sub-nationally in pilot areas in each of the three countries, by the EPI programme. Within the pilot area in each country, districts (in Ghana), subcounties (in Kenya), and clusters of equivalent population (in Malawi), have been randomized to either introduce the vaccine in 2019 (implementation clusters), or to delay introduction until the initial introduction has been evaluated (comparison clusters). Thus, the evaluation will use a cluster-randomized design. Community reporters will document all deaths occurring in children under 5 years of age in the implementation and comparison clusters. Hospital surveillance is being established in sentinel hospitals in a subset of clusters to monitor the incidence of admission with malaria and other conditions. Vaccine administration will be delivered through, and monitored by, the Expanded Programme for Immunization (EPI) in each country and uptake of RTS,S/AS01RTS,S will also be measured independently through cluster-sample household surveys.

The MVPE will determine whether the introduction of RTS,S/AS01RTS,S leads to a reduction in child deaths, and a reduction in the incidence of hospital admission with severe malaria, over the 4 years of the implementation programme (MVIP, Malaria Vaccine Implementation Programme). The MVPE also addresses three safety signals that were observed in the phase 3 trial but whose significance was unclear: an unexplained excess of meningitis cases in RTS,S/AS01 recipients, an excess in cerebral malaria cases among RTS,S/AS01 recipients who developed severe malaria and, among the relatively small number of deaths that occurred in the trial, an excess of deaths among girls who received RTS,S/AS01 compared to girls who did not receive RTS,S/AS01. The data to be collected

[1] In the phase 3 trial, for every 1000 children vaccinated with 4 doses, a total of 1774 (95% CI 1387–2186) cases of malaria, 40 (95% CI 19–64) hospital admissions due to malaria, and 19 (95% CI 4 to 35) cases of severe malaria, were averted over about 4 years.

over 4 years will inform a policy decision about wider use of RTS,S/AS01 as a means to reduce malaria morbidity and mortality, when provided in addition to existing malaria control measures. It is anticipated that sufficient data to assess the safety signals are likely to be available after the first 2 years of the evaluation. At that time, all available data would be reviewed, including outcomes related to safety of the vaccine in routine use (meningitis, cerebral malaria, mortality by gender, AEFI), the impact on severe malaria and all-cause mortality, coverage of the first 3 doses, preliminary data on administration of the 4th dose, and cost-effectiveness assessment. A recommendation about wider use of the vaccine might therefore be made before final impact data are available, recognizing that the impact in any setting will depend on coverage and timeliness of vaccination and the malaria burden, which may differ from those in the MVIP [3].

2. Design of the evaluation

A total of 158 clusters, 66 in Ghana, 46 in Malawi and 46 in western Kenya, each cluster with a total population of about 100,000 on average, were randomized to either introduce RTS,S/AS01 vaccine in 2019 or to delay introduction until a decision is reached about safety and effectiveness.

Throughout the evaluation areas, surveillance will be maintained to record all deaths in children under the age of 5 years. This will be done by community reporters who will record date of death, date of birth, age at death, gender, and residence location. The reporters will notify project staff who will visit the home of each child who has died to complete a verbal autopsy (VA). VAs will be completed for all deaths under 5 years of age. The purpose of the VA is to confirm key details (that the death occurred, the data of death and age at death, gender, where the child was living at the time of death and their normal place of residence if different), and to establish the likely cause of death. Causes of death will be determined from the VAs by three independent pysicians or computer algorithm. An important purpose of assigning causes of death (in children in vaccine-eligible age groups) will be to be able to exclude deaths due to injury from the analysis of effects of the malaria vaccine on overall mortality. At the time of the verbal autopsy (and if possible from the card by the local reporter) vaccination status will also be recorded, copying details from the home-based record (HBR) and completing a questionnaire to elicit caregiver recall about vaccinations. . In some circumstances, a limited VA will be performed instead of a full VA, collecting only the key details, without determining cause of death, other than determining death was due to accident or injury.

Surveillance for severe illness, with a focus on meningitis and severe malaria, will be maintained in part of the evaluation area in each country, through 18 sentinel hospitals, 8 in Ghana, 4 in Malawi and 6 in Kenya. These hospitals draw patients from a subset of clusters. The combined catchment areas include 32 of the clusters in Ghana, 17 clusters in Malawi, and 28 clusters in Kenya. Details of all inpatients aged 1 to 59 months will be captured, including residence location at the time of admission and their normal place of residence if different, date of admission, age at admission, date of birth, gender, final diagnosis, outcome (died or discharged alive), and clinical and laboratory details (for suspected cases of meningitis and severe malaria). Vaccination status will also be recorded for all admissions, copying details from the home-based record (HBR) or completing a questionnaire to elicit caregiver recall about vaccinations.

Table 1: Number of clusters randomized in each country, number of surveillance hospitals, and the number of clusters within the catchment of the sentinel hospitals

Country	Clusters	Population: annual birth cohort mean (range)	Sentinel Hospitals	Clusters in sentinel hospital catchments	Catchment clusters: annual birth cohort mean (range)
Ghana	66	3975 (912,8954)	8	32	4628 (1202,8954)
Kenya	46	5488 (2702,10739)	6	28	5559 (3487,10739)
Malawi	46	4820 (2816,8112)	4	17	5115 (2816,8112)
TOTAL	158	4661	18	77	5074

(excludes areas where the GSK phase 4 study is being done in each country)

Hospital and mortality surveillance will start when RTS,S/AS01 vaccination begins, or as soon as possible after that date, and be maintained until 46 months from the start of RTS,S/AS01 vaccination.

Administration of doses of RTS,S/AS01 and of other vaccines will be recorded by the EPI programme in each country using their normal record keeping system. Vaccination coverage will also be measured independently though community surveys. A survey will be undertaken at baseline, to measure coverage of EPI vaccines and Vitamin A, deworming treatment, malnutrition by MUAC, and the prevalence of *P.falciparum* infection (using an HRP2 Rapid Diagnostic Test). A second survey will be undertaken after about 18 months to measure coverage of EPI vaccines and, in RTS,S clusters, coverage of 1,2 and 3 doses of RTS,S in children 12-23 months of age. A third survey is planned to measure coverage of the fourth dose of RTS,S.

The MVPE was designed to detect differences in disease and mortality rates, over 46 months, between RTS,S and comparison clusters. Specifically, it was powered to detect a reduction in all-cause mortality in each country in RTS,S clusters of 10% of more (with 90% power) ; to detect a 2.6-fold or greater increase in the rate of meningitis (all countries combined) with 90% power ; to detect an increased incidence of cerebral malaria if the true rate is increased by 1.7-fold or more; and to detect an increase in the female:male mortality ratio with 90% power if the ratio is increased by 1.15-fold or more.

3. Description of the clusters

In Ghana, the clusters are districts, immunization clinic catchments fall entirely within single districts. In Malawi, clusters do not correspond to existing administrative units (no suitable existing units could be identified). Clusters were defined by allocating 187 health facilities that administer childhood vaccinations, into 46 groups of adjacent facilities, the combined catchment areas of each group having approximately the total population required for the evaluation. In Kenya, clusters correspond to subcounties.

4. Randomization

A separate document describes the randomization in detail [4]. The EPI programme in each country led the process of randomization and vaccine introduction. WHO provided technical assistance to assure unbiased randomization options. In each country the Ministry of Health held a public event to select the final randomization option. Randomization options were prepared using constrained randomization, which aimed to ensure implementation and comparator cluters were similar with respect to factors related to malaria burden, access to care and to immunisation services, whilst retaining a substantial element of chance in the allocation of each cluster. This is important so that the process remains fair and the observations in each cluster can be considered independent, an assumption for normal methods of statistical analysis to remain valid.

Randomization in each country was limited to permutations for which implementation and comparator areas would have clusters of similar population size, EPI coverage, and malaria transmission intensity would be similar in, and such that there would be a similar number of health facilities [4]. In addition, in Kenya and in Malawi, there was the geographic constraint that each county (in the case of Kenya) and each district (in Malawi) should have at least one cluster allocated to implement the RTS,S vaccine. Lastly, clusters containing a sentinel hospital were to be equally divided between RTS,S and comparator arms. A further constraint, that a measure of the number of hospital admissions from each cluster in the catchment of each hospital should be balanced between RTS,S and comparator clusters, was considered and was applied to some hospitals but not to all as this would have resulted in an overly constrained design. Catchment areas as defined, are also served by other facilities, thus the sentinel hospitals capture a proportion of hospital admissions in the defined areas.

3. Vaccination

In Ghana and Kenya , in RTS,S clusters, children receive RTS,S at 6, 7 and 9 months of age and the fourth dose at 24 months. All countries recommend monthly well child checks and weighing until 5 years of age, but attendance at these monthly visits tends to drop off during the first year of life. Vitamin A is given from 6 months of age at 6 monthly intervals and deworming from 12 months of age. The 6, 7, 9, 24 month schedule leverages the 6 month vitamin A visit, 9 month measles/rubella and/or yellow fever visit and 24 month vitamin A and deworming visit.

In Malawi , in RTS,S clusters, children should receive RTS,S at 5, 6 and 7 months of age and the fourth dose at 22 months. The 5, 6, 7, 22 month schedule leverages the 6 month vitamin A visit.

RTS,S vaccination started on 23rd April 2019 in Malawi, 30th April 2019 in Ghana, and 13th September 2019 in Kenya (Figure 1). There was no catch-up vaccination in Ghana or Malawi. Children born

before November 1st 2018 (in Malawi and Ghana) are not eligible for RTS,S vaccination. In Kenya, children were eligible for their first dose of malaria vaccine if they were aged 6-12 months.

Figure 1: Evaluation timelines

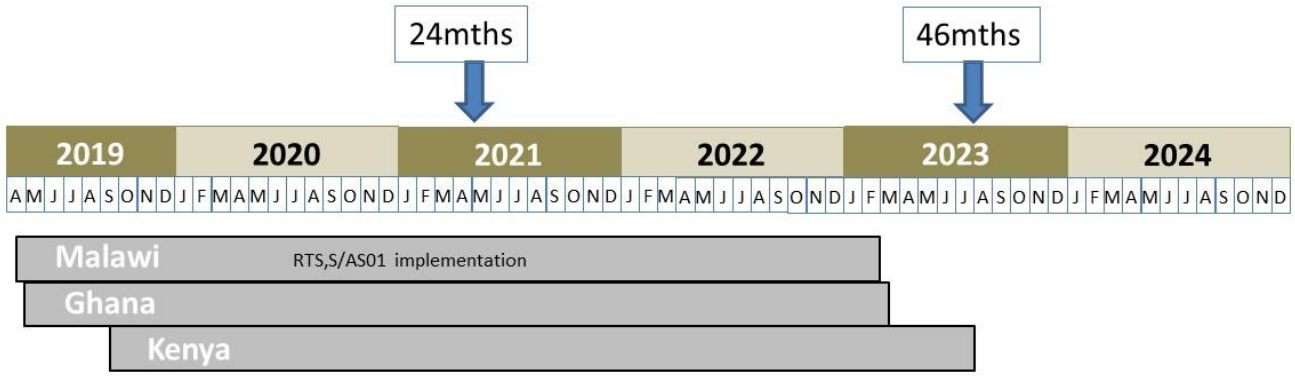


Table 2: EPI schedules

Integration of RTS,S/AS01 into the childhood vaccination schedule

World Health Organization

Child Age / Vaccine/1	Birth	6 wks	10 wks	14 wks	5 mo	6 mo	7 mo	9 mo	12 mo	18 mo	22 mo	24 mo
BCG	1											
Oral polio	0	1	2	3								
DTP-HepB-Hib (penta)		1	2	3								
Pneumococcal conj.		1	2	3								
Rotavirus		1	2									
Inactivated Polio				1								
Meningococcal A conj.										1		
Measles-Rubella								1		2		
Yellow Fever (Ghana, Malawi only)								1				
RTS,S in Ghana						1	2	3				4
RTS,S in Kenya						1	2	3				4
RTS,S in Malawi					1	2	3				4	
Vitamin A						1			2	3		4
Growth Monitoring	●	●	●	●	●	●	●	●	●	●	●	●
Deworming												●

4. Objectives

The main objectives related to assessment of impact, safety and feasibility are:

Primary objectives

Impact

to estimate the effect of RTS,S/AS01 introduction on

- all-cause mortality (excluding deaths due to injuries),
- the incidence of hospital admission with severe malaria, overall and in each country.

Safety

- to estimate the effect of RTS,S introduction on the incidence of hospital admission with meningitis, (data pooled across the three countries)
- to estimate the effect of RTS,S introduction on the incidence of hospital admission with cerebral malaria (data pooled across the three countries)
- to estimate the effect of RTS,S introduction on all-cause mortality in boys and girls and to determine whether there is any evidence that RTS,S increases mortality in girls, overall and in each country
- to describe the frequency and profile of RTS,S/AS01 reported AEFI

Feasibility

- to estimate the proportion of children aged 12-23 months in RTS,S clusters who received three doses of RTS,S/AS01 by 12 months of age (in each country)
- to estimate the proportion of children aged 27-38 months in RTS,S clusters who received their fourth dose of RTS,S/AS01 by 27 months of age (in each country)

Secondary objectives

Impact

to estimate, overall and in each country, the effect of routine delivery of RTS,S/AS01 on:

- incidence of hospital admission with severe malaria anaemia
- incidence of hospital admission with cerebral malaria (by country)
- incidence of hospital admission, by specific cause
-
- incidence of hospital admission for any cause
- incidence of hospital admission for any cause except malaria
- incidence of hospital admission with blood transfusion (or the requirement for transfusion)
- incidence of malaria-specific mortality in hospital (from hospital diagnosis)
- incidence of hospital deaths due to malaria, in boys and girls

Safety

Explore the association between RTS,S/AS01 and AESI, as agreed with each country's immunization program and regulatory authority, and with the Data Safety Monitoring Board

Feasibility

In each country:

- to estimate the coverage of recommended EPI vaccines in children from areas implementing RTS,S/AS01 and in children from areas not implementing RTS,S/AS01
- to estimate the proportions of children receiving each individual dose (the first, second, third, fourth, as appropriate) for each recommended vaccine
- to estimate the coverage and utilization of ITN/LLIN, IRS and any other recommended malaria prevention and control measures, in children from areas implementing RTS,S/AS01 and in children from areas not implementing RTS,S/AS01
- to document patterns of health-seeking behaviour for febrile children among children from areas implementing RTS,S/AS01 and in children from areas not implementing RTS,S/AS01
- to assess if the introduction of additional contacts between 5-9 months of age alters dropout rates for routine vaccinations and changes the number of fully vaccinated children
- to assess whether the introduction of RTS,S/AS01 is associated with a change in the coverage of other key childhood interventions, including anti-helminth administration (deworming) and Vitamin A supplementation

5. Outcomes

The primary outcomes for assessing impact are mortality from all causes, excluding deaths due to injury, and hospital admission with severe malaria.

The primary outcomes for assessing safety are hospital admission with probable meningitis, hospital admission with cerebral malaria, and mortality (excluding deaths due to injury) by gender.

The primary outcomes for assessing feasibility of achieving high coverage of RTS,S, are receipt of three doses before 12 months of age, and receipt of a fourth dose before 27 months of age.

Primary and secondary outcomes are defined in Table 3 and in the protocol [1].

Table 3: Definition of evaluation outcomes

Outcome	Definition
Mortality (excluding deaths due to injuries)	Death due to any cause, except injury
Mortality (all causes)	Death due to any cause
Malaria-associated mortality in hospital	Death in hospital in an individual who has confirmed malaria
Severe malaria	*P.falciparum* infection detected by RDT (antigenaemia detected by RDT or in the absence of RDT result, parasitaemia by microscopy at any density) AND one or more of the following: a) impaired consciousness (Glasgow coma score<11, Blantyre coma score<3, or assessed as P or U on the AVPU score and *not positive* for probable or confirmed meningitis; b) multiple of atypical convulsions (more than two episodes within 24 hours or prolonged (>15minutes), or focal) and *not positive* for probable or confirmed meningitis; c) respiratory distress (manifested as chest indrawing or deep breathing); d) severe malaria anaemia (haemoglobin concentration <5g/dL or haematocrit <15%).
Cerebral malaria (broader)	*P.falciparum infection* (by RDT or microscopy, as above) with impaired consciousness (Glasgow coma score <11 or Blantyre coma score <3 or assessed as P or U on AVPU score), excluding cases with CSF findings consistent with probable meningitis.
Cerebral malaria (stricter, LP required)	*P.falciparum infection* (by RDT or microscopy, as above) with impaired consciousness (Glasgow coma score <11 or Blantyre coma score <3 or assessed as P or U on AVPU score) AND CSF findings not consistent with probable meningitis (this strict cerebral malaria diagnosis cannot be made without an LP).
Severe malaria (stricter)	*P.falciparum* infection (as above) AND one or more of the following: a) impaired consciousness (Glasgow coma score<11, Blantyre coma score<3, or assessed as P or U on the AVPU score and negative for meningitis (LP performed); b) multiple of atypical convulsions (more than two episodes within 24 hours or prolonged (>15minutes), or focal) and negative for meningitis (LP performed); c) respiratory distress (manifested as chest indrawing or deep breathing); d) severe malaria anaemia (haemoglobin concentration <5g/dL or haematocrit <15%).
Probable meningitis	Children will be considered to have “probable meningitis” if in a suspected case*, the macroscopic aspect of the CSF is turbid, cloudy or purulent; or the CSF leukocyte count is >10 cells/mm3.
Confirmed meningitis	Any suspected or probable case, laboratory confirmed by culture or PCR to be of bacterial, viral or other aetiology in the CSF.
Malaria-associated anaemia	haemoglobin <11g/dL (or PCV<33%) and positive for *P.falciparum* (as above) in a hospital patient
Severe malaria-associated anaemia	haemoglobin <5 g/dL (or PCV <15%) and positive for *P.falciparum* (as above) (in a hospital patient)
Transfusion	Blood transfusion ordered or provided in a hospitalized child positive for *P.falciparum* (as above)
Hospital admission (any cause)	A stay in hospital/inpatient facility for at least one night, (or patients who were admitted but died before an overnight stay was completed).
Hospital admission (malaria)	Hospital admission with *Plasmodium* antigenemia detected by RDT (or in the absence of RDT result, parasitaemia by microscopy at any density)
Hospital admission (not malaria)	Hospital admission without a diagnosis of malaria
Febrile convulsion	Adapted from Brighton case definition for generalised seizures:

	Generalised seizures that occur in a febrile# child (6–60 months old) who does not have intracranial infection, metabolic disturbance or history of afebrile seizures.

#Fever defined as axillary temperature >=37.5ºC unless considered otherwise by country guidelines.

*Suspected meningitis: A child with one or more of the following present (with or without fever): neck stiffness, two or more seizures in the last 24 hours, bulging fontanelle, convulsions (partial, complex febrile or other atypical presentations), seizures if less than 6 months or greater than 6 years , altered consciousness (Blantyre Coma Score <3 or, Glasgow Coma Score <11 or P or U on the AVPU scale [Alert Verbal Painful Unresponsiveness Scale]) or any other clinical symptoms indicative of meningitis or cerebral malaria by clinical judgement. Lumbar puncture will be encouraged in all such children, according to national diagnostic and treatment guidelines, for examination of cerebrospinal fluid (CSF).

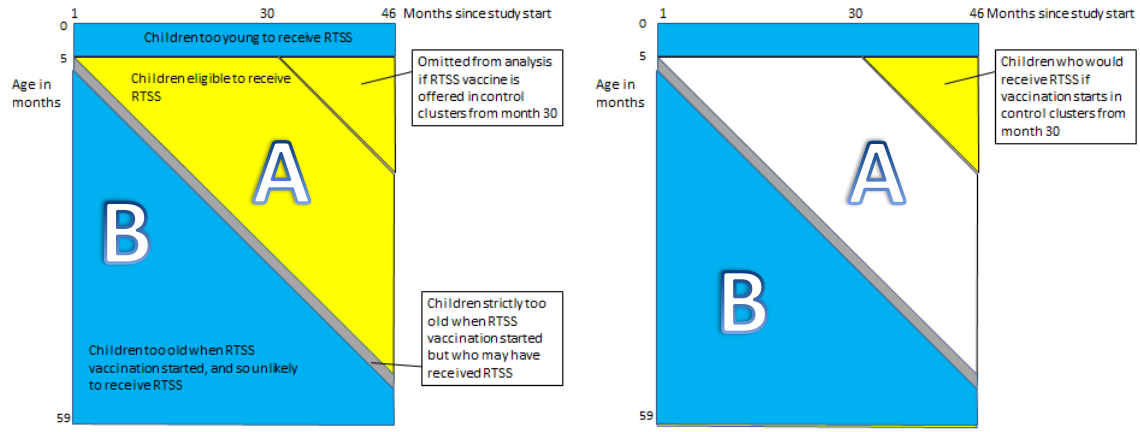
Alternative case definitions for some outcomes will be explored in secondary/sensitivity analyses.

6. Analysis approach

In each RTS,S cluster, children under 5 years of age can be classified into those eligible to receive RTS,S and those who are not eligible either because they are too young or there were too old when RTS,S vaccination started. Classification into eligible (group A) and non-eligible (group B) is based on date of birth and age, not on whether they actually received RTS,S, so the classification can be applied in exactly the same way in both intervention and comparator clusters (Figure 2):

Figure 2. Classification of age-eligible and non-eligible children in the RTS,S vaccination and comparator areas (Malawi example).

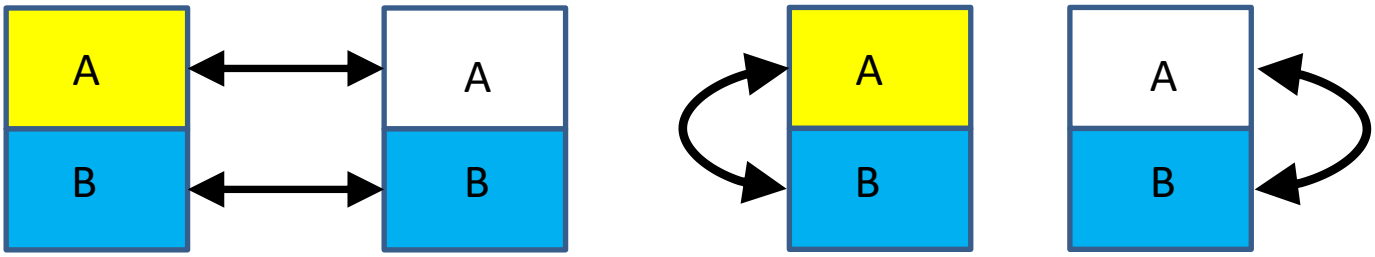
Intervention clusters (Malawi): Comparison clusters (Malawi):



Assuming introduction of RTS,S has no influence on incidence in non-eligible children, the difference in incidence between group B in the RTS,S arm and group B in the comparator area, reflects differences unrelated to RTS,S (such as differences in access to hospital, differences in underlying rates, or differences in diagnostic performance of hospitals or, for deaths, differences in the completeness of mortality surveillance). The difference in incidence between group A in the RTS,S arm and group A in the comparator arm, reflects these same differences *and* the effects of the introduction of RTS,S. We would like to subtract the Group B differences from the Group A

differences to get a better estimate of the effect of RTS,S introduction. Equivalently, we can consider the difference between Group A and Group B in the comparator area, and subtract this from the difference between Group A and Group B in the RTS,S arm (Figure 3); this way round is advantageous because the initial differencing is done within each cluster, thus eliminating the differences between the clusters in baseline incidence, access to hospital and diagnostic performance of the hospital etc. The test to assess the effect of RTS,S introduction amounts to a test for interaction between vaccine arm and group (A or B), [6,7 (p272-4)].

Figure 3: Comparison of between implementation and comparator areas, of the within-cluster differences.



The same approach can be used to determine if the effect of RTS,S on mortality differs between girls and boys. The mortality rate ratio between girls and boys (who are eligible for vaccination) in clusters in the RTS,S arm, is compared with the mortality rate ratio between eligible girls and boys in clusters in the comparator area.

This analysis approach takes account of potential imbalance between the implementation and comparator areas. In each country, surveillance for meningitis and severe malaria will be maintained in a subset of clusters, through selected hospitals. Each hospital draws patients primarily from a small number of clusters. These clusters may not be well balanced between intervention and comparator with respect to population size, the underlying risk of meningitis or malaria, access to hospital, and the diagnostic performance of the hospital. For mortality, although a larger number of clusters has been randomized, there is also a risk of imbalance if it is not possible to implement surveillance uniformly well in all clusters. Pre-intervention data on these outcomes is not available as surveillance was established for the purpose of the evaluation, at the same time as the vaccine was introduced.

Methods for estimating rate ratios and rate differences are detailed below.

Assumptions of the analysis approach:

The approach assumes that:

a) the incidence by age is such that there are expected to be enough cases for analysis (in both group B and group A)

b) the age of cases will be determined sufficiently accurately to assign them to group A or B reliably

c) there is limited catch-up vaccination, so that a group can be defined based on age who would not have received RTS,S vaccine

d) for rate ratios, we have assumed that the ratio of person time between Group A and group B is the same in each cluster and in both intervention and comparator areas although this assumption could be relaxed if cluster-specific data were available. The population size of the cluster does not need to be known to estimate rate ratios, but for rate differences, population denominators are needed. For both effect measures an assumption is that the ratio of incidence of the outcome of interest between eligible and non-eligible age groups, would be the same in implementation and comparison areas in the absence of the malaria vaccine. We would expect this to hold for other conditions, this could be checked by comparing incidence ratios of outcomes unrelated to malaria. Such "tracer" conditions could then be used to adjust for an imbalance, by comparing the ratio (eligible:non-eligible) for the outcome of interest to the corresponding ratio for the tracer conditions, and comparing this double ratio between implementation and comparator areas.

e) for analysis of meningitis: the causative pathogen could differ by age, but we ignore the pathogen in this analysis, we are looking for a possible enhancement of the risk of meningitis independent of the causative organism. Meningitis can occur in outbreaks, if an outbreak or outbreaks occurred and were highly localised, and if age groups were affected differently, this could make results difficult to interpret, and a different approach may be needed (e.g. comparing numbers of vaccinated and unvaccinated cases within the outbreak with the percentage vaccinated in the population or ideally, using a case control study).

f) an implicit assumption of this approach is that indirect effects (herd immunity effect, i.e. effect on malaria transmission, an associated beneficial "community effect" reducing malaria among those not vaccinated) are not important. Such indirect effects of RTS,S on transmission are not captured, they cancel out in this analysis (as groups A and B benefit equally from these indirect effects). It is not anticipated that RTS,S would have any appreciable indirect effects so this may not be considered a limitation in this evaluation. (Clearly this is not relevant for the safety endpoints.) (Note: the method does capture 'indirect effects' in the sense of a greater reduction in mortality or severe disease than would be expected based on prevention of malaria alone, due to the contribution of malaria to severity of diseases other than malaria such as malnutrition, pneumonia, non-typhoidal salmonella).

g) Effect of RTS,S on coverage of other vaccines: The introduction of RTS,S could potentially improve coverage of other vaccines, if this had an effect on outcomes e.g. meningitis, this wouldn't matter in the sense that the analysis estimates the net effect of the RTS,S programme.

7. Analysis populations

For each death and each hospital patient, cluster of membership will be determined from the location of normal residence. Within each implementation and comparator cluster, the following groups can be defined:

A: children who would be eligible for RTS,S vaccination

A1: Children who would be eligible to have had the first dose of RTS,S according to their age and date of birth (in Ghana born on or after Nov 1st 2018 and aged at least 6 months, in Malawi born on or after November 1st 2018 and aged at least 5 months , in Kenya born on or after September 1 2018 and at least 6 months of age). This is shown for the case of Malawi, as group A1 in Figure 4 and 5.

Three subsets of this group are defined as follows:

A2: Children who would be eligible to have received 3 doses of RTS,S according to their age and date of birth (born on or after November 1st 2018 and aged at least 8 months in Malawi and born on or after Nov 1 2018 and aged at least 10 months in Ghana and born on or after Sep 1 2018 and aged at least 9 months in Kenya) (A2 in the Figure). It is likely that 3 doses are needed for maximum protection, this group therefore excludes children not old enough to have received their third dose.

A3: Children who would be eligible to have received 3 doses of RTS,S according to their age and date of birth (as above) AND who would be eligible for their first vaccine dose during the first 12 months of the project, and are aged less than 40 months (A3 in the Figure). This group are of interest because the age distribution reflects the age distribution in the population, whereas in groups A1 and A2, younger ages are over-represented. Impact in this group therefore better reflects the impact that might be expected in the longer run.

A4: The protocol specifies for some analyses, restricting events to children who have received 3 doses of DTP-Hib-HepB. The MVPE is designed to estimate the overall, public health impact of introducing RTS,S vaccine. The impact observed will depend on the level of coverage of RTS,S and timeliness of doses. Limiting analysis to children who received three doses of pentavalent vaccine aims to more closely approximate the impact in RTS,S recipients as the coverage of RTS,S is expected to be higher among those that received penta3 than in the general population. For meningitis, there is also a concern that many of the cases may occur in children who are under-vaccinated, and therefore if RTS,S did increase the risk of meningitis, this effect may be obscured. In this analysis, any cases who had not had penta3 would be excluded (even if they had received RTS,S).

This group is not strictly comparable between implementation and comparator areas, as the additional immunization contacts for RTS,S provide extra opportunities to receive pentavalent vaccine which are not available in the comparator areas. If surveys or administrative data indicate a difference in penta3 uptake between areas, the potential for bias will need to be borne in mind when interpreting the results for this indicator.
. (DTP3 coverage in DHS surveys in Ghana, Malawi and Kenya, has been high, the coverage of DTP3 in Ghana was 85% in the 2014 DHS survey [8], in Malawi 93% (DHS 2015-16, [9]) and Kenya 90% (DHS 2014, [10]). If penta3 coverage is similar in the implementation and comparator areas, limiting analysis to DTP3 recipients may not greatly alter estimates of impact).

B: Children who would not be eligible for malaria vaccination

These are an important comparison group within each cluster. Children aged under 5 months (Malawi) or under 6 months (Ghana, Kenya), should not receive RTS,S. Children who were too old when RTS,S immunisation started, that is children born before Oct 24 2018 (in Malawi), 1 Oct 2018 (in Ghana), should not receive RTS,S vaccine. Children born just before the cut-off date may receive RTS,S but children born two months prior to these dates, are unlikely to have received RTS,S. These are marked B in the Figure.

In Kenya, children will be able to receive their first dose up to 12 months of age.

Primary analysis of safety will be based on events in group A1 (children eligible for RTS,S). Primary analysis of impact will be based on events in group A2 (children eligible for RTS,S and old enough to have received 3 doses). Secondary analyses of safety and of impact will be done limited to recipients of 3 doses of pentavalent vaccine (group A4).

Figure 4: Definition of analysis populations in Malawi, according to age in months and calendar time from month 1 (May 2019) to month 46

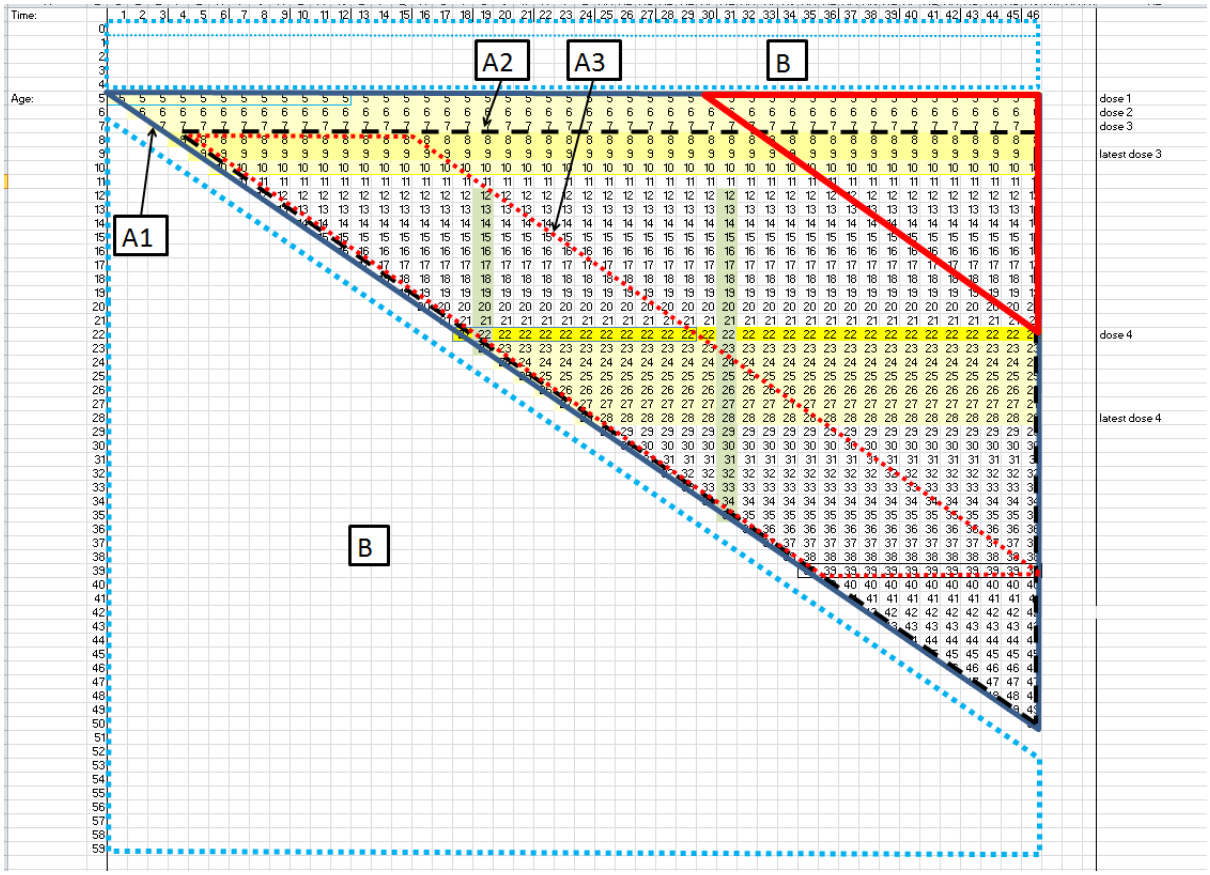
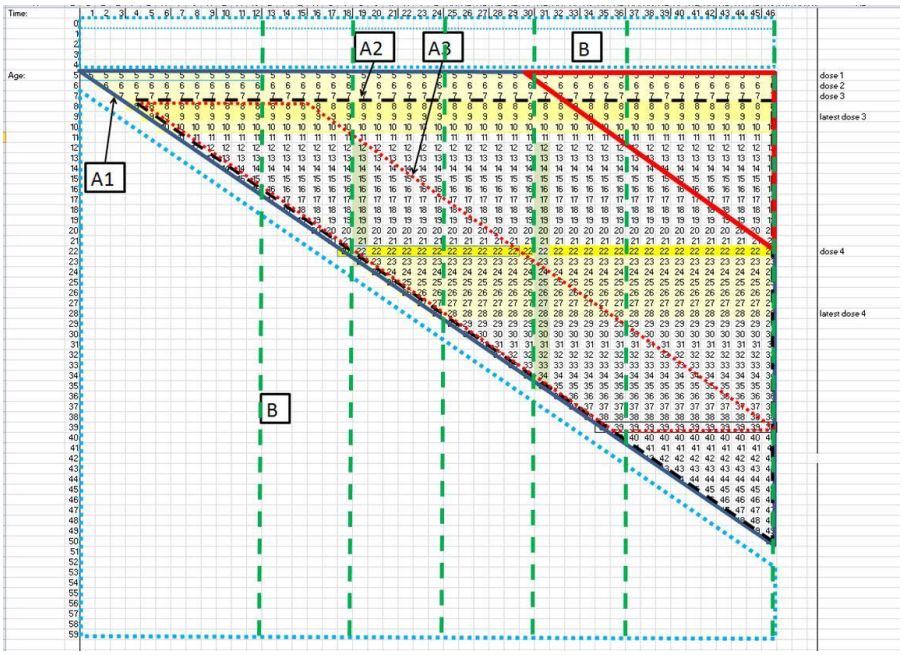


Figure 5: Populations included in analyses at 12,18,24,30,36 and 46 months



All deaths in children 1-59 months and hospital admissions 1-59 months will be recorded and the events in non-eligible age groups will be included in analyses.

A1, A2 and A3 are groups which based on their age would be eligible for RTS,S vaccination. Children who would not be eligible for malaria vaccination, because they are too young or because they were too old when the programme started, are an important comparison group within each cluster. These groups are marked B. Children born just before the cut-off date may receive RTS,S but it is assumed that children born before September 2018 (in Malawi and Ghana) or January 2019 (Kenya), who would have been 8 months old when RTS,S vaccination started, are unlikely to have received RTS,S.

Groups are defined in the same way in intervention and comparator clusters.

The vertical light green bars show the age range for coverage surveys, envisaged to take place at month 19, and at month 31. A survey at month 19 would measure coverage of 3 doses of RTS,S in children aged 12-23 months who received their first dose of vaccine during the first 12 months of RTS,S introduction. A survey at month 31, would measure coverage of 3 doses of RTS,S in children aged 12-23 months, who should have received their first dose during the second year of the programme, and the coverage of four doses in children aged 28-35 months who should have received dose 4 18-25 months after the start of the programme (the first children to receive dose 4).

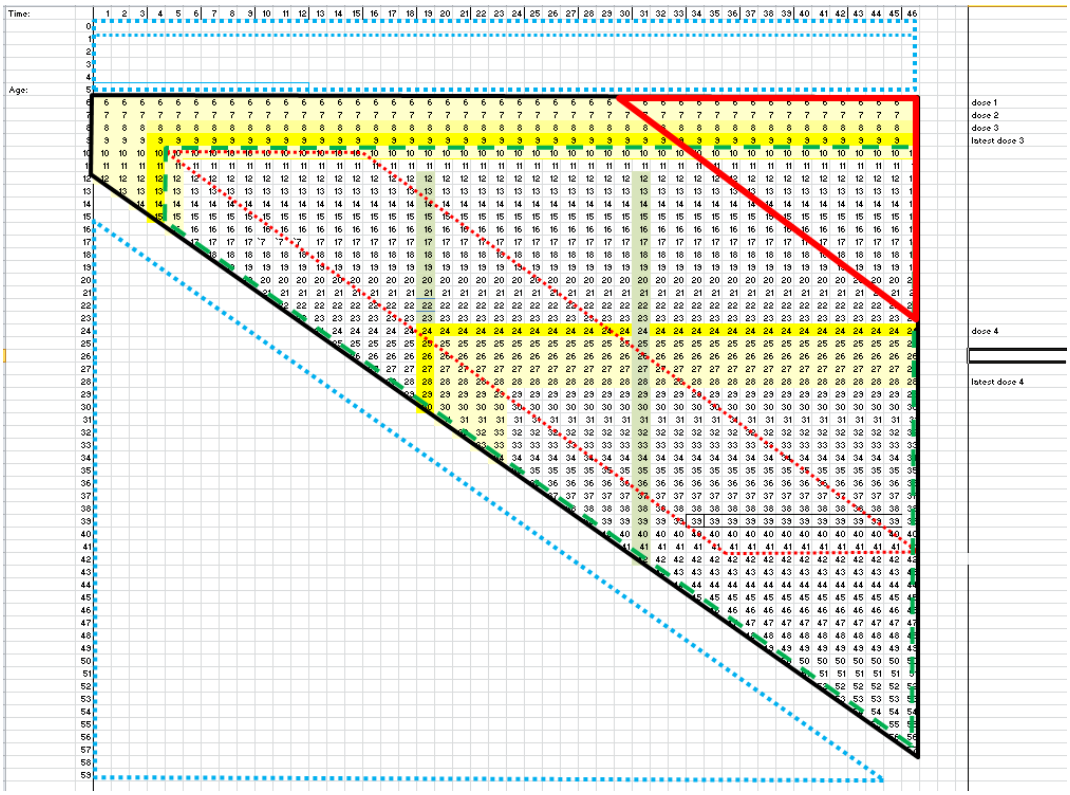
If vaccination were to start in comparator clusters before the end of the evaluation, the corresponding upper right triangle (shown here in red for vaccination starting in comparator clusters in month 30) would be excluded, in both intervention and comparator clusters.

The horizontal blue dashed line at 1 month of age indicates that for analysis of effects of RTS,S on mortality, we will exclude neonatal deaths from group B, as the factors influencing variation in neonatal mortality between clusters may be different from those influencing variations in post-neonatal mortality (and there may be a relatively large number of neonatal deaths). Including these in the within-cluster controls, may therefore not be appropriate.

The layout in Ghana is similar but with vaccination starting at 6 months of age, and dose 3 at 9 months of age.

In Kenya, the layout is different as it is planned to have catch-up vaccination up to 12 months of age (Figure 6).

Figure 6 Definition of analysis populations in Kenya



8. The effect of coverage, contamination and confounding on expected effect sizes; power to detect or exclude effects of a given size; and the timing of analyses

Dilution of effect sizes due to incomplete vaccine coverage, contamination, and confounding:

The analysis compares rates in implementation and comparator clusters. The observed difference (rate difference or rate ratio) reflects the effect of the vaccine in vaccinated children, and the proportion of person time spent vaccinated. Indirect effects (impact of the vaccine on malaria transmission) are unlikely to be important. If there is no contamination, the rate ratio comparing intervention:comparator areas is:

$R = \{ [c \times \Sigma n_v/\Sigma T_v + (1-c) \times \Sigma n_u/\Sigma T_u] \div [\Sigma n'_u/\Sigma T'_u] \} \times$ [correction term], where:

Σn_v and ΣT_v are the total events and total person time, respectively, in vaccinated children in vaccine-eligible age groups in implementation clusters;

Σn_u and ΣT_u are the corresponding totals in unvaccinated children in vaccine-eligible age groups, in implementation clusters;

$\Sigma n'_u$ and $\Sigma T'_u$ are the corresponding unvaccinated totals in vaccine eligible age groups in comparator clusters (all assumed unvaccinated)

$c=\Sigma T_v/(\Sigma T_v+\Sigma T_u)$ is the overall proportion of vaccinated person time, in vaccine-eligible age groups in intervention clusters.

The [correction term] is $[\Sigma m/\Sigma Y] / [\Sigma m'/\Sigma Y']$, this is a correction factor which uses data for age groups that are mnot eligible for RTS,S/AS01 vaccine, to adjust for randomization imbalance; Σm and ΣY are the total events and total person time in non-eligible age groups in implementation clusters and $\Sigma m'$ and $\Sigma Y'$ the corresponding values in comparator clusters (calculation of R explained section 20).

Effect of confounding: The overall rate ratio can be written as: $R = c.R_v+(1-c)R_u$, where R_v is the rate ratio comparing vaccinated children in implementation areas and unvaccinated children in comparator areas, and R_u is the rate ratio comparing unvaccinated children in implementation and comparator areas, and R is the overall rate ratio that we observe directly. If there is no confounding (as if vaccination essentially assigned at random), $R_u=1$ and $R=cR_v+1-c$. The observed R could then be compared with expected value if R_v was equal to the rate ratio in vaccinated children in the phase 3 trial for a given value of c. If there is confounding, $R_v=(FxR_v^*)$, $R_u=(GxR_u^*)$, where R_v^* is the true (unconfounded) rate ratio vaccinated:unvaccinated, $R_u^*=1$ by definition, and F and G represent the confounding, i.e. F is the factor by which the rate in the RTSS-vaccinated is increased relative to the unvaccinated comparator population, aside from any effect of the RTSS vaccine, and G the

factor by which the rate in those that did not receive RTSS, is increased relative to the comparator population.

Confounding with respect to malaria outcomes: Preliminary estimates from the baseline surveys in each country, during which children 5-48 months were tested for malaria infection using an RDT and receipt of routine vaccinations was recorded, there was no evidence that coverage of basic vaccines was associated with malaria prevalence, suggesting confounding with respect to malaria outcomes may not be important.

Confounding with respect to meningitis: Apart of any effect of RTSS itself, the rate of meningitis is likely to be higher in the RTSS-unvaccinated and lower in the RTSS-vaccinated, because children who receive RTSS also more likely to have had pneumococcal and Hib vaccines, which protect against meningitis, than children who did not receive RTSS, unless high levels of coverage of pneumococcal and Hib vaccines have essentially eliminated vaccine serotypes from circulation. Conservatively, if say 90% of the population received anti-meningitis vaccines, and coverage of RTSS-1 is 80%, and 84% of those who received anti-meningitis vaccines also receive RTSS-1, then coverage of anti-meningitis vaccine among those who receive RTSS-1 is 95% and 70% in those who do not receive RTSS-1:

		RTSS-1		
		Y	N	Total
Anti-meningitis vaccine	Y	760	140	900
	N	40	60	100
	Total:	800	200	1000
Coverage of pneumo-2:		760/800=0.95	140/200=0.7	900/1000=0.9

If efficacy of of anti-meningitis vaccine against meningitis of any cause is 70%, then the rate in RTSS recipients (before any RTSS effect) is reduced by a factor 0.95x(1-0.7)+(1-0.95)= 0.335, compared to a factor of 0.9x(1-.07)+0.1=0.37 in the unvaccinated comparator population, a rate ratio of 0.335/0.37=0.905. Among those who do not receive RTSS-1, the corresponding factor is 0.7x(1-0.7)+0.3=0.51, compared to 0.37 in the comparator population, a rate ratio of 0.51/0.37=1.378:

	RTSS-1	**no RTSS-1**
anti-meningitis vaccine coverage:	0.95	0.7
rate of meningitis reduced by factor:	0.95x(1-0.7)+0.05=0.335	0.7x(1-0.7)+0.3=0.51
Vomparator population:		
anti-meningitis vaccine coverage:	0.9	
rate of meningitis reduced by factor:	0.9x(1-0.7)+0.1=0.37	
rate ratio:	0.335/0.37=0.905	0.51/0.37=1.378

The observed rate ratio comparing implementation and comparator areas is: $R=c.(FxR_v^*)+(1-c).G$, with F=0.905, G=1.378 in this example. If the true (unconfounded) rate ratio was $R_v^*=10.5$ (as in the phase 3), and coverage of RTSS-1 was 80%, we would expect to see a rate ratio R=0.8x0.905x10.5+(1-0.8)x1.378=7.878 between implementation and comparator areas. To be able to rule out an association with meningitis of the magnitude seen in the phase 3 trial we would therefore want to be able to exclude rate ratios of about 7.9 or more. This compares with a value of 0.8x10.5+0.2=8.6 if there was no confounding.

Contamination:

Contamination could occur due to:

a) Children in comparator clusters receive the malaria vaccine (either by attending a clinic in an implementation area instead of "their own" clinic, OR by moving address, after they have been vaccinated, from an implementation to a comparator cluster), so there is some benefit (or harm) in comparator areas, diluting our estimates of impact which compare implementing and comparator clusters;
b) Children in intervention clusters receive their vaccinations from a clinic in a comparator area (by going to the "wrong" clinic, or by moving address after completing vaccination), so coverage of the vaccine in intervention areas, is lowered;
c) Cases or deaths, who are assigned to the wrong cluster by mistake.
d) Cases or deaths are assigned to the wrong age group, i.e. cases in vaccine-eligible children assigned to non-eligible age groups and cases in non-eligible age groups assigned to the vaccine eligible age group, due to error in age or date of event. And then there is also contamination whereby children outside the eligible age group, receive the vaccine.

If there is contamination due to having some vaccinated children in age-eligible groups in comparator clusters, and the proportion of vaccinated person time among vaccine eligible age groups in comparator clusters is d,

$$R = \{ [c \times \Sigma n_v/\Sigma T_v + (1-c) \times \Sigma n_u/\Sigma T_u] \div [d\Sigma n'_v/\Sigma T'_v + (1-d)\Sigma n'_u/\Sigma T'_u] \} \times [\text{correction term}],$$

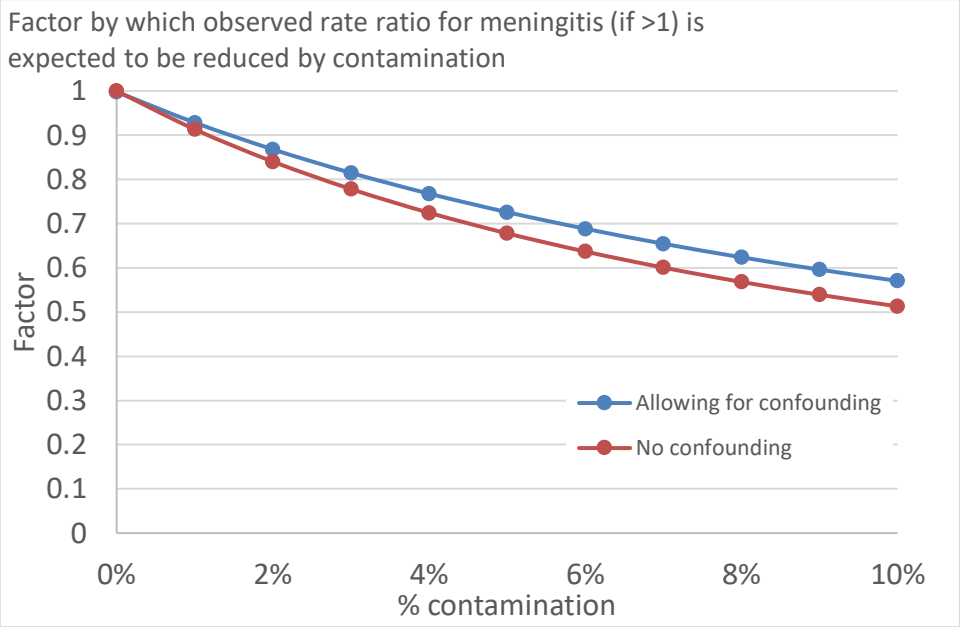
where $\Sigma n'_v$ and $\Sigma T'_v$ are the total events and person time in vaccinated children in vaccine-eligible ages in comparator clusters. Comparing this with the corresponding R without contamination, the estimate with contamination differs by a factor: $1/(d.R_v^*F'+(1-d).G')$, where Rv* is the unconfounded rate ratio for vaccinated:unvaccinated, and F' and G' represent the confounding. If the coverage of RTSS-1 in comparator areas is 1%,

		RTSS-1		
		Y	N	Total
anti-meningitis vaccine	Y	10	890	900
	N	0	100	100
	Total:	10	990	1000
coverage of anti-meningitis vaccine:		10/10=1	890/990=0.89	900/1000=0.9

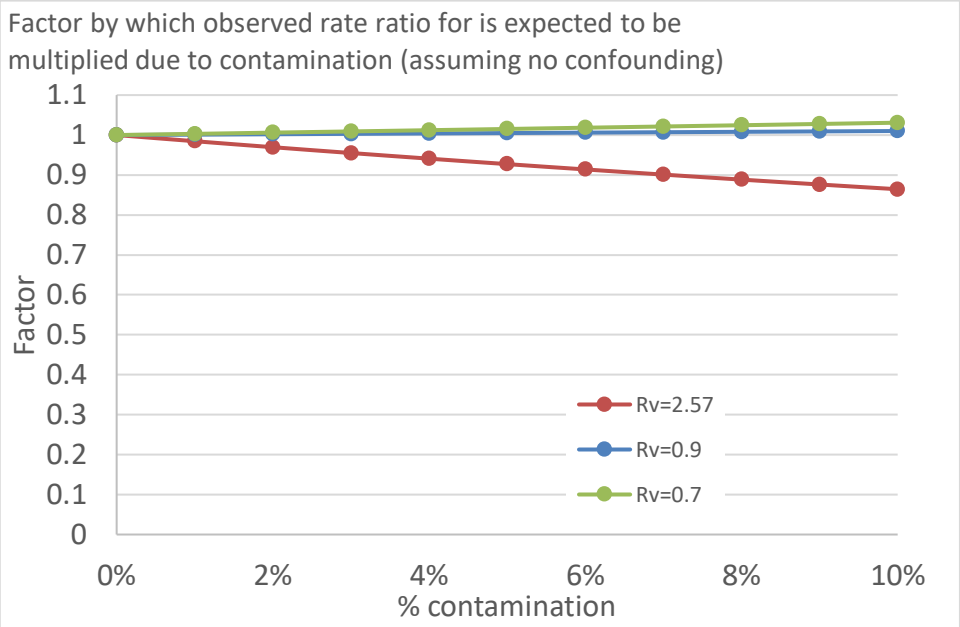
	RTSS-1	**no RTSS-1**

anti-meningitis vaccine coverage:	1	0.89
rate of meningitis reduced by factor:	1x(1-0.7)=0.3	0.89x(1-0.7)+(1-0.89)=0.37
Comparator population:		
anti-meningitis vaccine coverage:	0.9	
rate of meningitis reduced by factor:	0.9x(1-0.7)+0.1=0.37	
rate ratio:	0.3/0.37=0.81	0.37/0.37=1

So for 90% coverage of anti-meningitis vaccine, and 1% coverage of RTSS, if all the RTSS-vaccinated received anti-meningitis vaccine, F'=0.81 and G'=1. If Rv* is 10.5, and d is 1%, this gives R with contamination reduced by a factor of 1/[0.01x10.5x0.81+(1-0.01)x1]=0.93. If d is 2% this factor is 0.87 and 5%, 0.73. The values with no confounding are 0.91,0.84,0.68 for 1%, 2% and 5% contamination. So allowing for confounding in implementation areas, and contamination and confounding in comparator areas, if contamination is 1%, we would aim to exclude a target R value of 7.878x0.93=7.3, or for d=2% or 5%, values of 6.8 or 5.7.



Effect of contamination on other outcomes (examples for cerebral malaria (Rv* 2.57); mortality, 0.9, severe malaria, 0.7):



Power calculations:

Power was estimated using simulations in which the number of events in cluster j and age group k, was a random value from a poisson distribution with mean equal to $r.k_{ij}y_{jkt}\theta_{jk}$, where r is the assumed underlying incidence rate, k_{ij} is an adjustment factor for the relative access to hospital *i* for cluster *j* (explained below), y_{jkt} is the person time in cluster *j* in age group *k* at time *t*, and θ_{jk} is the rate ratio for the outcome associated with vaccine introduction, for cluster *j* and age group *k*. $\theta_{jk} = \theta$ for vaccine-eligible age group in clusters in the intervention arm, $\theta_{jk} = 1$ otherwise. For each simulation, the ratio of the number of events between the two age groups was compared between arms using a ratio estimator. The simulation was repeated 10,000 times, to determine the distribution of the estimate of θ and of the 95% confidence limits. Simulations were done for null value (θ =1) for a range of assumed underlying incidence rates, and for various values of θ >1 (safety) and θ <1 (impact) to estimate power to detect or exclude effects of interest. Simulations were repeated for a range of values of the underlying incidence rate, for various time points, in order to determine the number of events that would be required to have adequate power, at each time point.

Relative access to hospital:

k_{ij} represents the relative access to hospital *i* from cluster *j*, k_{ij}=1 for the cluster in which the hospital is located, 0<k_{ij}<1 for other clusters in the catchment area, and k_{ij}=0 for other clusters. *k* values were estimated for each hospital, using data on the number of admissions under 5 yrs from each cluster in the catchment area, for a one-year period before intervention started: $k_{ij}=(a_{ij}/n_j)/(a_i*/n*)$, a_{ij} = admissions to hospital *i* from cluster *j*, n_j=population in cluster *j*; a_i* is the number of admissions to hospital *i* from the cluster the hospital *i* is located in, and $n*$ is the population of that cluster. If

access in all clusters in the catchment was the same as for the cluster the hospital is located in, (and the reason for lower admissions in outlying clusters was purely access to hospital or the choice of hospital, rather than a difference in underlying incidence in the communities), then the number of admissions would be increased by a factor: $\sum_{i,j}(a_{ij}/k_{ij}) \,/\, \sum_{i,j} a_{ij}$. These factors were 1.6 for Ghana, 1.0 for Malawi, and 2.2 for Kenya, based on the data for the pre-intervention period for children under 5 yrs of age. These factors could be recalculated using data during the evaluation, by using data for the non-eligible age groups in each cluster, for all cause admissions or for particular outcomes, and used to assist in interpretation of rates and rate differences.

Power to detect or exclude effects of a given size:

For safety outcomes, the question posed in the WHO malaria vaccine position paper was whether the excess of cases of meningitis and cerebral malaria observed in the phase 3 trial, which were unexplained, were causally related to the vaccine. We therefore estimated the number of events required for 90% power to exclude or detect rate ratios of this magnitude, after allowing for dilution due to vaccine coverage being less than 100%, and due to confounding and contamination.

For impact outcomes, we estimated the number of events (cases of severe malaria, or deaths due to any cause excluding injury) to have 90% power to detect an impact consistent with the effects observed in the phase 3 trial, after allowing for the level of coverage and contamination and, in the case of deaths, the likely proportion of deaths caused by malaria.

Table 4: Administrative estimates of average coverage of RTS,S/AS01-1 in implementation areas

(average for the period from vaccine introduction to Dec 2020), and estimates of the degree of contamination (% admissions (or % deaths) in vaccine-eligible age groups in comparator areas, who had received RTSS-1).

		Contamination (RTSS-1 coverage in comparator areas in hospital catchments)	
Country	Coverage of RTSS-1	from mortality surveillance (HBR or recall)	from hospital admissions (HBR or recall)
Malawi	72.3%	4.8%	2.7%
Ghana	69.0%	0.0%	5.9%
Kenya	79.8%	5.3%	4.6%
TOTAL	73.7%	3.4%	4.4%

Table 5: Power to detect or exclude effects on safety outcomes, of the magnitude observed in the phase 3 trial, and power to detect impact on hospital admission with severe malaria, and on all-cause mortality

	Effect in the phase3 trial[1]	Corresponding effect size, allowing for coverage[2]	Effect size allowing for contamination[3]	Effect size allowing for confounding[4]	Number of events required for given power to detect or exclude effects of the given size[5]	
Safety outcomes:					90% power	80% power
Meningitis	10.5	8.00	5.80	5.66	90	70
Cerebral malaria	2.57	2.16	2.03		400	300
Gender interaction (mortality)	1.7	1.52	1.47		2000	1500
Impact:						
Severe malaria[6]	0.66	0.75	0.76		4000	3000
Mortality[7] a)	0.898	0.92	0.93		24000	18000
b)	0.864	0.90	0.90		17600	13100

1 In the phase 3 trial, the observed rate ratio for meningitis (RTS,S/AS01:control) was 10.5, and for cerebral malaria, 2.57. Among girls, there was a higher mortality rate in RTS,S/AS01 recipients than controls, the mortality ratio was 2 over the whole trial period, and was 1.7 in the first part of the trial (from enrolment until the 4th dose was administered) and 3.4 in the latter part of the trial.

2 Assuming coverage of RTSS-1 of 74% (the average value in Table 4)

3 Assuming coverage of RTSS-1 in comparator areas is 4%

4 Confounding: For meningitis, assuming RTSS-recipients more likely to have received other vaccines that would prevent meningitis (90% coverage of anti-meningitis vaccines, 50% efficacy against all-cause meningitis, all RTSS recipients have received anti-meningitis vaccine). Based on preliminary data from baseline surveys, EPI coverage did not appear to differ between high and low malaria prevalence areas, and was similar in children who tested positive for malaria infection by RDT and children who tested negative. We have therefore assumed no confounding for impact outcomes.

5 The number of events required for given power. The numbers refer to the total events in eligible and non-eligible groups combined, except for the interaction by gender, where the number refers to the number of deaths in vaccine-eligible ages.

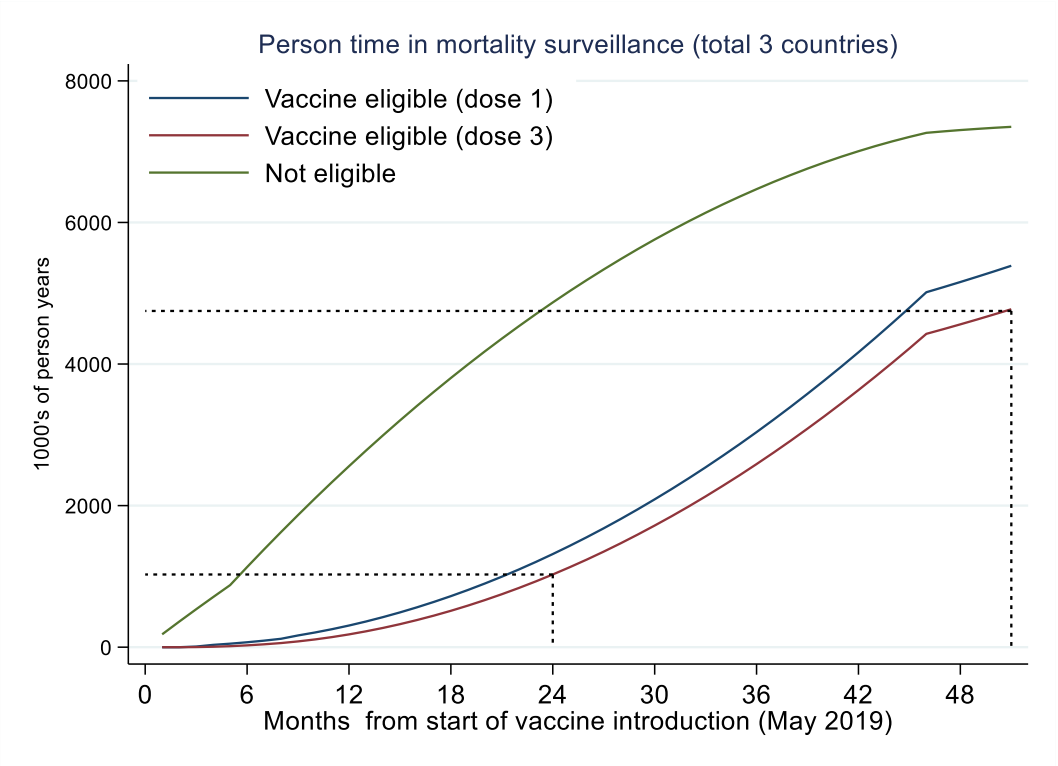
6 Efficacy of RTS,S/AS01 against severe malaria, during the first 20 months of the trial (ITT analysis) was 34% (rate ratio 0.66).

7 The expected reduction in all-cause mortality is a) 10.2%, if efficacy against malaria deaths is 34%, and the percentage of deaths caused by malaria (in the absence of the vaccine) is 30%. This translates to a 7% reduction after allowing for 74% coverage and 4% contamination. If the percentage of deaths caused by malaria is 40% (b), the expected reduction in all-cause deaths is 14%, which translates to a 10% reduction at population level.

Further details of power simulations, including predicted distribution of point estimates of effects, are given in Annex 4. Based on incidence rates of the key outcomes observed during the first 12 months of surveillance, we estimated that by the end of month 24 (April 2021), sufficient events would be expected to have accrued to give 90% power to detect or exclude a 5-fold increase in incidence of hospital admission with meningitis associated with introduction of RTS,S/AS01, and at least 80% power to detect a 2-fold increase in incidence of hospital admission with cerebral malaria, and 90% power to detect a 1.4-fold relative increase in all-cause mortality in girls compared to boys associated with RTSS/AS01 introduction, in pooled analysis across the three countries. By this time point, there would also be good power to detect a reduction in the incidence of hospital admission with severe malaria. There would be low power to detect reductions in all-cause mortality at this time point, but it is anticipated that sufficient events would have accrued by the end of the evaluation (month 46 in each country) to have 90% power to detect a reduction of 10% and 80% power to detect a reduction of 8%.

Data will be reviewed by the DSMB at 6-monthly intervals. If there is evidence of increased incidence of safety outcomes in implementation clusters in vaccine-eligible age groups, vaccination status of cases in these groups would be examined, to determine if the proportion vaccinated is consistent with an excess given the level of vaccine coverage from administrative data; sensitivity analysis would be done to explore effects on results of uncertainties in age, missing data, and alternative case definitions; incidence of all cause admissions and admissions for conditions unrelated to the vaccine, would be compared between implementation and comparator areas to understand potential biases; and analysis will be done to determine if the results are similar in all three countries. If after completing these analyses, the data are consistent with an increased risk of adverse outcomes associated with RTSS introduction, there will be an assessment of the magnitude of the excess risk in relation to evidence on potential benefits, to decide whether to recommend suspending or discontinuing vaccination.

Final analyses will be done when data up to month 46 are complete in all countries. Figure shows the accrual of person time, in the three countries combined, during the evaluation. During the latter half of the evaluation the total person time increases by a factor of more than 4.


Person time in mortality surveillance (total 3 countries)
Vaccine eligible (dose 1)
Vaccine eligible (dose 3)
Not eligible
1000's of person years
8000
6000
4000
2000
0
0
6
12
18
24
30
36
42
48
Months from start of vaccine introduction (May 2019)

9. Rebound effects

In the phase 3 trial, in children who received only three doses of RTS,S, the initial protection from severe malaria afforded by the vaccine was followed by a period in which their risk was increased compared to unvaccinated controls. This was thought to be a result of a reduced rate of acquisition of natural immunity in RTS,S recipients which led to an increased risk of severe malaria when the protection from the vaccine waned. The resulting excess of severe malaria cases offset the number of severe cases that had been averted during the previous months such that by the end of the trial there was no overall benefit in terms of severe malaria among children who received only 3 doses. In contrast, children who received four vaccine doses had sustained protection from severe malaria. It therefore appeared from the phase 3 trial data that a fourth vaccine dose was necessary in order to reduce a child's overall risk of severe malaria over the 4 years of the evaluation, although there was some statistical uncertainty surrounding this conclusion [2,5]. Beyond 4 years, based on data from three of the Phase 3 trial sites where surveillance was maintained for a total of seven years, it appeared that the risk of severe malaria is very low [3]. The implication is that national immunisation programmes would need to ensure children receive the fourth dose of the vaccine in order to reduce the burden of severe malaria and any associated mortality but, as far as severe malaria is concerned, a fifth dose would be unlikely to be necessary.

For these reasons (the questions about the feasibility of achieving high coverage with the 4-dose schedule, and about the importance of receiving the fourth dose) there is some uncertainty about the likely impact of the malaria vaccine in practice and an important aim of the MVPE is therefore to determine whether the introduction of RTS,S leads to a reduction in all-cause mortality over the 4 years of the project, and a reduction in the incidence of severe malaria, and to measure uptake of four doses of vaccine. The evaluation is not designed to measure the impact of the 4th dose specifically, but is designed to determine the overall public health impact of RTS,S introduction. If, over the 46 months of the evaluation, a reduction in mortality was observed in vaccine-eligible children, or there was a reduction in severe malaria with consistent findings for mortality, this would indicate that any rebound effects did not prevent vaccine introduction having a net beneficial impact, especially if consistent findings are seen in the subgroup of children who have, or could have, complete follow-up (i.e. up to at least 39 months of age, group A3 in Figure 2).

10. Denominators

Malawi: Estimates for 2021 of the total population, population under 5 years of age, and the number of live births, and of surviving infants, in the catchment area of each facility were provided by the EPI. These were based on estimates of the total population in the catchment of each facility from house to house enumeration by HSAs, which were rescaled so that the total for each district equalled the census estimate of the total population for the district for 2021. The number of children under 5, the number of livebirths and other population estimates were then obtained by applying the ratio to total population for the same year and district, from census projections. We summed the estimates for the facilities to obtain estimates for each MVIP cluster for 2021. We then applied estimates of the relative change in the number of live births and the population under 5, for each district, from census projections, for the yeas 2018-2023, and applied these factors to the cluster totals to obtain estimates of the number of births and the population under 5 in each year. We then

and applied census estimates of the ratio of girls to boys among under 5's in each district in each year, and census estimate of the sex ratio at birth, to obtain estimates of births and population under 5 for boys and girls in each cluster in each year.

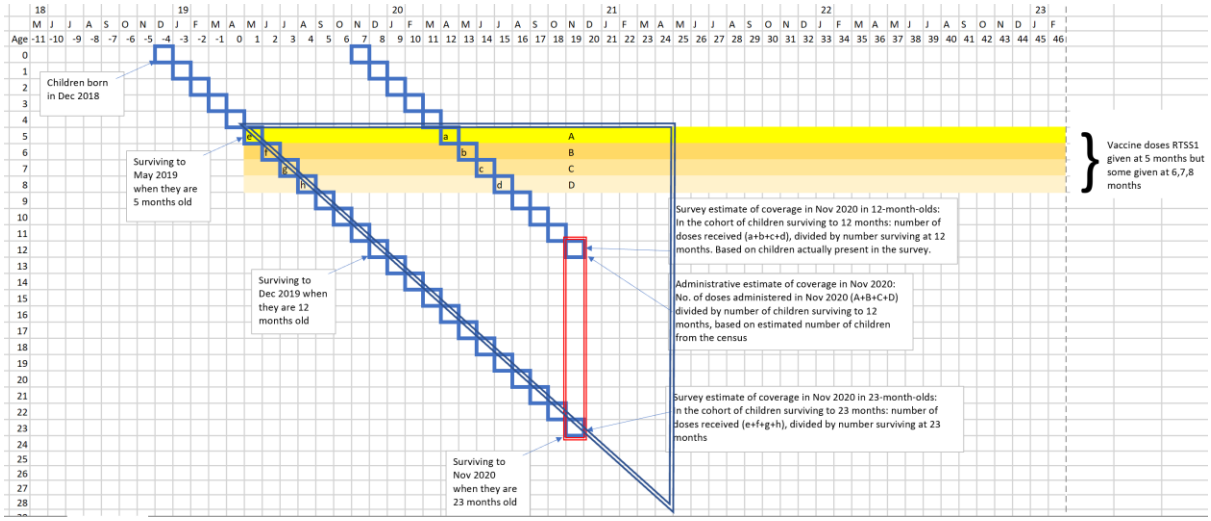
To estimate the population under 5 years of age, in each calendar month from Jan 2018 to Dec 2023, by month of age, in each MVIP cluster, an iterative method was used, by applying age-specific survival estimates, while preserving the census estimates of total annual births and mid-year totals under 5 years of age. We first estimated the number of births in each calendar month, and the total population under 5 in each calendar month, and then used an iterative procedure consisting of two steps, firstly applying age-specific survival to the population in each month to give the population in the next month, and then re-scaling to maintain the correct total population under 5, these steps being repeated until convergence. Census estimates of the infant and under5 mortality for the appropriate district [6], and the UN IGME estimate of the neonatal mortality rate [7] were used. This procedure gives cluster populations with annual mean under-5 total, and annual births, equal to census estimates, with age structure consistent with the census estimates of infant and child mortality. The estimated population in each month of age for each calendar month were then used to calculate totals for vaccine-eligible and non-eligible children. These were used for calculation rates and rate differences, but were not used in estimation of rate ratios.

Ghana and Kenya: Similar approach will be used for Ghana and Kenya. In Malawi, population estimates, projections, and mortality rates were available by district from the 2018 census. In Ghana, census projections were available from the 2014 census and estimates of mortality rates from the 2014 DHS. In Kenya, provisional population estimates from the 2019 census were available, and estimates of mortality rates were obtained from the 2014 DHS.

Denominators for admininistrative coverage: Children under 5 years of age and especially under 1 year of age tend to be under-enumerated in national censuses [11,12]. It is therefore recommended in the WHO guidelines on estimation of denominators for immunization programmes [13], that population denominators for immunization should be determined from census estimates of the number of births per year, adjusted for reporting inaccuracies, and adjusted for survival. The reporting inaccuracies that are corrected for, relate mainly to the precision of recall of birth dates as to whether the birth was in the last 12 months before the census. As national estimates of birth rates are more reliable than sub-national estimates, it is common to use subnational estimates of population or other indicators such as (if coverage is high) the number of BCG doses administered, to allocate the estimated national number of births proportionately to subnational areas. The number of births each year is obtained using the national estimate of the annual growth rate in the number of births determined from population projections, by applying the age-specific birth rates to the projected population of women of reproductive age. We require estimates of the number of births each year from 2015 to 2023, in order to estimate the population under 5 years of age from 2019 to 2023. We will then calculate the number of births in each month from these annual totals, so they reflect steady growth month to month. Adjustment for survival to each time point , can be based on infant and children mortality rates estimated from the census, from DSS data on survivorship from DSS sites in each country, or from the age distribution of children in the baseline household surveys, or a combination of these.

These denominators can be checked by comparing administrative estimates of vaccine coverage with survey estimates. A scaling factor can then be inferred for the denominator (no. of surviving infants), to bring survey and admin estimates of coverage into agreement. The relationship between survey and administrative estimates of coverage is illustrated in Figure 8.

Figure 8: Relationship between survey and administrative estimates of vaccine coverage



11. Cluster membership

Cluster of membership will be defined based on the normal place of residence. (Caregivers may have a temporary address, if they moved to be near hospital when their child was admitted; cluster of membership will be determined based on their normal address).

It is questionable to reassign cases in comparator clusters who received RTS,S, to an adjacent RTS,S cluster, as this would add events to RTS,S clusters, but with no assignments in the opposite direction from RTS,S to comparator (children in an RTS,S cluster who missed RTS,S because they went to a clinic in a comparator cluster for vaccination). Using vaccination status to assign to clusters could therefore lead to a bias.

Potential for contamination will be assessed from estimates of RTSS coverage in comparator areas and in non-eligible age groups, in the midline surveys, and these estimates used to assess the impact of contamination on effect sizes.

For hospital outcomes, hospital cases coming from outside the defined catchment clusters (as per the randomization) will be excluded from analysis.

12. Assignment of events to analysis groups

Each event (death, case of meningitis, cerebral malaria, severe malaria, etc) will be assigned to a vaccine-eligible or non-eligible group based on their age and date of birth. Events which fall in the grey area (not strictly eligible but could receive RTS,S) will be excluded. Age will be determined by calculation from the date of birth and the date of death or hospital admission, and/or from the declared age in months. Where age is not known to the nearest month, age will be imputed and sensitivity analysis used to explore the effects of this (for example if the age is known in integer

years, the analysis will be repeated, assigning the age to a random age within the 12-month interval, and then means of the resulting log rate ratios and standard errors will be reported).

Indicators of the quality of data on exact age at death or admission to hospital: For each main outcome, the proportion with age determined from a record of the date of birth; the proportion from caregiver recall of the data of birth; the proportion of dates of birth with missing month, the proportion with missing day. Normal date formats do not allow for missing day or month, it is important that dates are captured in such a way that it is recorded if the day is missing or if the month is missing. It is desirable to ask about the age of the child as well as the date of birth.

13. Completeness of surveillance

Completeness of surveillance will be monitored during the project and will be reported at each analysis time point. Indicators of completeness are,

For mortality:

- the number of weekly reports per village or per reporter, tabulated per month per cluster in each arm
- in each area of the evaluation: the number of deaths reported, and the number of these with VA completed, and the time interval between death and VA, by cluster and eligibility group and by gender

For meningitis:

- for each hospital in each month, the number of patients admitted in the target age range, the number eligible for lumbar puncture, and the number who had lumbar puncture performed with CSF results, tabulated by gender and vaccine eligibility group.

For all types of event, overall and by cluster and by eligibility-group:

- % events with missing information on age, address, vaccine status
- % events with incomplete information on age, address, vaccine status
- % events with vaccine status by HBR, % by recall only, % missing
- % of children whose details were reported by the primary caregiver and the % reported by a proxy for the caregiver
- Check for duplicated events: number of events of each type with the same location, gender, age and date.

As a check on completeness of mortality surveillance, midline household surveys may include questions about under 5 deaths in the last year, and check to see if those deaths were captured by the surveillance system.

14. Laboratory quality control

For meningitis, CSF samples will be sent to ….for identification of causative organism by PCR. (Note that there is likely to be a delay, samples to be sent in shipments quarterly and analysis likely to take 4-6 weeks. However, confirmed meningitis is not a primary outcome).

15. Vaccination status

Indicators of data quality

for each of the main outcomes, the proportion of events where vaccination status was determined from the home-based record (HBR); the proportion of these where the vaccination status was cross-checked against caregiver recall; the proportion of events where a home-based record was not available and of these the proportion where vaccination status was obtained by caregiver recall. For each event type, for the proportion where vaccination status was cross-checked against clinic registers (for those with an HR and those without an HBR). These indicators will be compared between the areas of the evaluation and between children eligible for RTS,S (group A) and not eligible (group B).

Results of validation studies

will be reported, where these have been done – comparing, in a sample of children who have an HBR, their vaccination status according to caregiver recall, HBR, and from clinic registers, and in a sample of children who do not have an HBR, comparing vaccination status from caregiver recall and the clinic register.

Table 6: Dummy table for comparison of different sources of vaccination status

		No. RTS,S doses received					
Children without an HBR:							
	Clinic register:	0	1	2	3	4	missing
Recall:	0						
	1						
	2						
	3						
	4						
	Don't know						
Children with an HBR:							
	Clinic register:	0	1	2	3	4	missing
HBR:	0						
	1						
	2						
	3						
	4						
	Clinic register:	0	1	2	3	4	missing
Recall:	0						
	1						
	2						
	3						
	4						
	Don't know						
	HBR:	0	1	2	3	4	
Recall:	0						
	1						
	2						
	3						
	4						
	Don't know						

Comparison between groups

For each main outcome, the proportion of cases/admissions who received 0,1,2 3 and 4 doses of RTS,S will be tabulated by cluster, eligibility group, and implementation/comparator area. This will indicate the degree of contamination (RTS,S recipients in comparator clusters), and whether children in group B in RTS,S clusters, may have received RTS,S. Among deaths which occurred in children of vaccine-eligible age in RTS,S/AS01 areas, the proportion who had received RTS,S/AS01 should be similar in boys and girls, if there is no differential effect of the vaccine on survival of boys and girls, since uptake of vaccination is generally very similar for boys and girls and this would be expected to hold also for RTS,S/AS01. Among cases of cerebral malaria, the proportion who received RTS,S/AS01 would be expected to by similar to the RTS,S/AS01 coverage in the general population of the same age, if RTS,S/AS01 has no effect on incidence of cerebral malaria. For meningitis, the comparison of vaccination status of cases with vaccine coverage in the population is complicated by the fact that receipt of RTS,S/AS01 is likely to be strongly associated with receipt of pentavalent and pneumococcal vaccine, which could mask an association with RTS,S/AS01.

Table 7: Dummy table for comparison of vaccination status of cases between areas

Outcome (4 tables: all deaths; all admissions; meningitis; severe malaria)	Vaccine	RTS,S Group A	RTS,S Group B	Comparator Group A	Comparator Group B
	RTS,S doses 0	% (n/N)			
	1				
	2				
	3				
	4				
	pentavalent doses 0				
	1				
	2				
	3				

16. EPI data

Checks for completeness and consistency will be performed as recommended in the WHO Guidance for immunization programme managers [16], using disaggregated (facility-level) data in DHIS2. All three countries are using DHIS2 to manage their routine EPI data. The WHO Data Quality Tool, a DHIS2 application, can be used to review data quality and completeness. For each vaccine, outlier checks will be used to screen the monthly reports from each facility (values more than 3 standard deviations from the annual mean for the facility being flagged), and for repeated adjacent numbers, and zero values will be flagged for checking to confirm if the value should be treated as zero or missing. Consistency checks will include comparing the number of doses in a facility for vaccines that according to the schedule should be given together (e.g. penta1, OPV1, PCV1 and Rota1), and checks to detect when the number of successive doses increases (e.g. penta 2>penta1). Summary tables will present indicators of completeness and consistency by cluster and arm.

The number of doses of RTS,S1, RTS,S2, RTS,S3 and RTS,S4 (and each of the other vaccines) will be tabulated by month and area in each country, and for each cluster. Drop-out rates (%reduction from RTS,S1 to RTS,S2 etc.) will be calculated for each vaccine with multiple doses, including RTS,S, by area and by cluster and month. Estimates of number of doses administered as a percentage of the estimated target population (administrative 'coverage') will be calculated for each cluster, using denominators derived for the impact and safety analyses.

Drop-out rates and administrative 'coverage' for pentavalent vaccine and other vaccines will be compared between RTS,S and comparator areas.

More detailed analysis of EPI data related to RTS,S, will include impact of RTS,S delivery on the uptake and coverage of other vaccines. This will be through a comparison of antigen coverage, by dose, in vaccination areas compared with comparator areas.

17. Vaccination coverage (RTS,S and other vaccines)

Reporting of vaccination coverage surveys will follow recommendations in the 2018 WHO manual [17, chapter 6]. These will be representative sample household surveys of children 5-48 months of age for the baseline and endline surveys and children 12-23 months for the midline survey.

Baseline surveys

- the percentage of children aged 12-23 months who received each vaccine (and each dose), and the percentage fully vaccinated, with 95% confidence intervals. Calculation of sampling weights is described in Annex 2 for the baseline surveys in Ghana, Malawi and Kenya.
- cross-tabulation by implementation/comparator area, gender, socio-economic status, and by transmission intensity strata (strata based on the prevalence of *P.falciparum* by RDT).
- Kaplan Meier curves showing timing of vaccine doses (among those with HBR)
- prevalence of *P.falciparum* (positive RTD) by area

Midline surveys

Surveys conducted at approximately month 18, counting from the start of vaccination, will measure coverage of 3 doses of RTS,S in children aged 12-23 months who were eligible to receive their first dose in the early part of RTS,S introduction. A survey at approximately month 30 will measure coverage of 3 doses of RTS,S in children aged 12-23 months, who were eligible for their first dose during the second year of the programme, and the coverage of four doses in children aged 28-35 months who should have received dose 4 by 27 months of age (the first children to receive dose 4).

Estimates of % children 12-23 months who received 0,1,2 and 3 doses of RTS,S, and Kaplan Meier curves showing the timing of vaccine doses and the % receiving 3 doses by 12 months of age.

Estimates of the % children 28-35 months who received a fourth dose, and the % who received dose 4 by 27 months of age.

The influence of RTS,S introduction on coverage and timeliness of other vaccines

Coverage of penta 3, MR1 and (if relevant) yellow fever among children 12-23 months of age will be compared between RTS,S and comparator areas (taking account of the correlation between surveys when calculating confidence intervals if the same survey PSUs are used for baseline and midline surveys). Timing of these vaccine doses will also be compared between arms, using Kaplan Meier plots.

Order of vaccination, and co-administration of RTS,S with other vaccines

% of children who were fully vaccinated for their age, when they received RTS,S1;
% children who received RTS,S3 with YF or MR (Ghana);
% children who had received less than 2 doses of RTS,S when they received YF or MR (Ghana);
% children who had received less than 2 doses of RTS,S when they received YF or MR (Ghana) who did not receive RTS,S with YF or MR (missed opportunities);
% children who received RTS,S4 who had previously received less than 3 doses of RTS,S.

18. Other interventions that might affect MVPE outcomes

A diary of other interventions (e.g. LLIN distribution, IRS campaigns, azithromycin MDA, expansion of community case management, creation of new health facilities) maintained for each cluster would allow reporting of contextual information about the possible influence of other interventions taking place in study clusters during the MVIP. These are being recorded by the evaluation partners and reported to WHO as part of quarterly technical reports. Likewise, the PATH and WHO MVIP team members in each pilot country records similar events that can affect EPI delivery of the RTS,S vaccine.

19. Calculation of rate ratios and rate differences

Incidence of the primary safety outcomes (probable meningitis, confirmed meningitis, cerebral malaria) and impact outcomes (deaths, deaths excluding injurys, and severe malaria) will be compared between comparator and RTS,S areas, in each country and for the three countries combined. For mortality, the evaluation is powered to estimate effects within each country by the end of the 46 months. For meningitis, primary analysis will be pooled across the three countries.

Rate ratios and rate differences, with 95% and 99% confidence intervals, will be calculated as in the table below. Secondary safety and impact outcomes listed in Table 5 will be analysed in the same way.

Table 8: Dummy table for effect measures for safety and impact outcomes

Outcome	Analysis populations* primary (secondary)	Comparator N/PYAR rate	RTS,S N/PYAR rate	Rate ratio (95%CI)	Rate difference (95%CI)
Safety:					
Probable meningitis	A1 (A4)				
Confirmed meningitis	A1 (A4)				
Cerebral malaria	A1 (A4)				
Impact:					
Deaths	A2 (A4)				
Deaths (excl. injury)	A2 (A4)				
Severe malaria	A2 (A4)				

*A1: eligible for RTS,S; A2: eligible for 3 doses of RTS,S; A4: received 3 doses of pentavalent vaccine

(Note: For severe malaria, if a reduction in the number of admissions for all-cause febrile illness is observed that is commensurate with the reduction in severe malaria, this would give reassurance that the reduction in severe malaria was not caused simply by a reduction in the prevalence of parasitaemia without prevention of severe illness.)

Subgroup analyses

Gender

A test of interaction will be used to determine whether the effect of RTS,S on mortality differs by gender, separately in each country, and for all countries combined, using population A1 (children eligible for RTS,S). We will also investigate whether the interaction by gender differs according to age, by considering age-eligible children in age strata (under 18 months, and 18 months and above, and, as data accrues, above 2 yrs).

Table 9: Dummy table for mortality rates by gender

	Comparator			RTS,S			
Outcome:	Girls	Boys	Rate ratio	Girls	Boys	Rate ratio	Interaction P-value
Deaths	Rate (N/PYAR)	Rate (N/PYAR)		Rate (N/PYAR)	Rate (N/PYAR)		
Deaths (excl. injury)	Rate (N/PYAR)	Rate (N/PYAR)		Rate (N/PYAR)	Rate (N/PYAR)		

The sex ratio of mortality varies by age. Newborn boys tend to have higher mortality than newborn girls, this relative advantage of girls tending to decrease later in infancy. Infant mortality rates tend to be higher in boys than girls, the sex difference in child mortality rates being less marked [18,19,20 p.72]. Estimates of sex differentials in childhood mortality published by the United Nations Population Division, showed that the average male:female ratio of age-specific mortality rates, for sub-Saharan Africa, based on data from surveys and censuses conducted in the 2000s, was 1.18 in infants and 1.03 in children 1-4yrs [20]. These patterns can vary, for example in Malawi, Ashorn *et al.* [21] found a 2-fold greater mortality risk in 1-2-year-old boys than girls in a cohort study of infants born to a group of 795 pregnant women. The male:female sex ratio of mortality rates tends to increase, in both infants and in children, as overall under-5 mortality decreases [18]. In infants, this trend is attributed to the fact that as overall under-5 mortality declines, early infant deaths, where girls have the greatest advantage, form a larger proportion of infant deaths. In children 1-4 years, this trend is attributed to a smaller proportion of deaths due to infection and a larger proportion due to non-infectious causes which are more common in boys. These patterns can change if girls have poorer access to treatment or nutrition than boys. These sex differences in mortality may be less evident in this analysis as we will exclude deaths in early infancy and deaths due to injury, but age, and the overall level of under-5 mortality, could be potential confounders in the assessment of the effect of RTS,S on the sex ratio of mortality. To explore these effects, the sex ratio of mortality will be compared between age groups and between areas with higher and lower overall levels of under 5 mortality, in each arm of the trial. And analyses of interaction by gender will be done with adjustment for age and for overall level of under-5 mortality.

As secondary analyses, other key outcomes will be also be analysed by gender (severe malaria; cerebral malaria; all cause admission to hospital).

Cerebral malaria

In addition to the primary analysis of cerebral malaria, comparing incidence between intervention and comparator areas, we will compare effectiveness of vaccine introduction between cerebral malaria and other forms of severe malaria using a test of interaction. If cerebral cases tend to occur at older ages a lower effectiveness might be expected due to waning vaccine efficacy, to check for

this we will plot the age distribution of cerebral cases and non-cerebral severe cases in the vaccine-eligible age group, in each country.

Transmission intensity

In each country, clusters will be classified into higher and lower transmission strata, based on the prevalence of *P.falciparum* by RDT in the baseline surveys. For impact outcomes, to allow for potential imbalance in age distribution of severe malaria between intervention and comparator areas, we will adjust for transmission stratum, as a secondary analysis. Impact for the non-malaria-specific outcomes (all cause mortality, all cause admission) will also be estimated separately in each stratum and a test of interaction will be done.

20. Statistical methods for estimating rates, rate differences and rate ratios

The layout of the data from each cluster and calculation of rate ratios and rate differences is as set out in the table below. *n* is the number of events, for example $n_{0,1,A}$ is the number of events in group A in cluster 1 in the comparator group (group 0), and $T_{0,1,A}$ is the corresponding person time. The totals in each arm are denoted by $n_{1,A}$ and $n_{1,B}$ (arm 1 = RTS,S) and $n_{0,A}$ and $n_{0,B}$ (arm 0 = comparator). Stata code to estimate rate ratios and differences is given in Annex 2.

Table 10: Calculation of rate ratios and rate differences

Country	Intervention	Cluster	Group A (events, person time)	Group B (events, person time)	Rate ratio (A/B) in each area	Difference between log rate ratios	Rate difference (A-B) in each area	Difference between rate differences
Ghana	RTS,S	1	$n_{1,1,A}$ $T_{1,1,A}$	$n_{1,1,B}$ $T_{1,1,B}$				
	RTS,S	2	$n_{1,2,A}$ $T_{1,2,A}$	$n_{1,2,B}$ $T_{1,2,B}$				
	…	…	…	…				
		TOTAL	$n_{1,A}=\sum n_{1,i,A}$ $T_{1,A}=\sum T_{1,i,A}$	$n_{1,B}=\sum n_{1,i,B}$ $T_{1,B}=\sum T_{1,i,B}$	$R_1=(n_{1,A}/n_{1,B})x(T_{1,B}/T_{1,A})$		$d_1=n_{1,A}/T_{1,A} - n_{1,B}/T_{1,B}$	
	Comparator	1	$n_{0,1,A}$ $T_{0,1,A}$	$n_{0,1,B}$ $T_{0,1,B}$				
	Comparator	2	…	…				
	…	….	…	…				
		TOTAL	$n_{0,A}=\sum n_{0,i,A}$ $T_{0,A}=\sum T_{0,i,A}$	$n_{0,B}=\sum n_{0,i,B}$ $T_{0,B}=\sum T_{0,i,B}$	$R_0=(n_{0,A}/n_{0,B})x(T_{0,B}/T_{0,A})$	$\log(R_1/R_0)$	$d_0=n_{0,A}/T_{0,A} - n_{0,B}/T_{0,B}$	d_1-d_0
Malawi	RTS,S	1	…	…	…			
	RTS,S	2	…	…	…			
	…	…	…	…	…			
	Comparator	1	…	…	…			
	Comparator	2	…	…	…			
	…	…	…	…	…	$\log(R_1/R_0)$		d_1-d_0
Kenya	RTS,S	1	…	…	…			
	RTS,S	2	…	…	…			
	…	…	…	…	…			
	Comparator	1	…	…	…			
	Comparator	2	…	…	…			
	…	…	…	…	…	$\log(R_1/R_0)$		d_1-d_0

Rate differences

$r_{0,A}= n_{0,A}/T_{0,A}$ is the incidence rate in the vaccine-eligible group (group A) in the comparator area (in one country), and $r_{1,A}= n_{1,A}/T_{1,A}$ is the corresponding rate in the RTS,S area, and $r_{0,B}= n_{0,B}/T_{0,B}$ and $r_{1,B}= n_{1,B}/T_{1,B}$ are the estimated rates in the non-eligible group (group B) in each area. These are ratio estimates of the rates [22], which give equal weight to each person year of observation, and so in the analysis each cluster is weighted according to its size.

The rate difference (comparator minus RTS,S) is:

$RD= [r_{0,A} – r_{1,A}] – [r_{0,B} – r_{1,B}]$,

this is the difference between the rates in the vaccine-eligible groups in the comparator area and the RTS,S area, $[r_{0,A} – r_{1,A}]$, less a correction factor $[r_{0,B} – r_{1,B}]$ which is the corresponding difference between non-vaccine-eligible groups in the comparator and RTS,S area. If evaluation areas were perfectly balanced, this correction term would be zero.

The rate difference can be rewritten $RD= [r_{0,A} – r_{0,B}] – [r_{1,A} – r_{1,B}]$, i.e. the difference between group A and group B in the comparator area, minus the difference between group A and group B in the RTS,S area.
In each area of the evaluation, the difference between group A and group B is calculated within the same clusters, this reduces the between-cluster variability in this difference, due to the correlation between the rates in A and B. The variance of $[r_{0,A} – r_{0,B}]$ is given by $V(r_{0,A})+V(r_{0,B})-2\text{cov}(r_{0,A} ,r_{0,B})$, [23, p.181], where:

$$V(r_{0,A}) = \left(\frac{m_0}{(m_0-1)T_{0,A}^2}\right)\sum_{i=1}^{m_0}\left(n_{0,i,A} - r_{0,A}T_{0,i,A}\right)^2$$

$$V(r_{0,B}) = \left(\frac{m_0}{(m_0-1)T_{0,B}^2}\right)\sum_{i=1}^{m_0}\left(n_{0,i,B} - r_{0,B}T_{0,i,B}\right)^2$$

$$\text{cov}(r_{0,A}r_{0,B}) =$$

$$\frac{m_0}{(m_0-1)T_{0,A}T_{0,B}}\sum_{i=1}^{m_0}\left(n_{0,i,A}n_{0,i,B} - r_{0,A}n_{0,i,B}T_{0,i,A} - r_{0,B}n_{0,i,A}T_{0,i,B} + r_{0,A}r_{0,B}T_{0,i,A}T_{0,i,B}\right)$$

where m_0 is the number of clusters in the comparator area. The variance of $[r_{1,A} – r_{1,B}]$ is calculated similarly. The variance of the rate difference (comparatorl-RTS,S) is $V(RD)=V(r_{0,A} – r_{0,B})+ V(r_{1,A} – r_{1,B})$, and the 100(1-α)% confidence interval is $RD \pm t_{\alpha/2,m_0+m_1-2}\sqrt{V(RD)}$. Rate differences will be calculated separately for each country. The rate difference RD is *the number of cases averted (or added) as a result of RTS,S vaccine being introduced into an area, in the age groups of children eligible for the vaccine, expressed per 1000 child years (or other suitable units)*. The estimated values will be specific to the analysis population chosen (A1, eligible for RTS,S; A2 eligible for 3 doses of RTS,S; or A3 eligible for 3 doses and followed up to 39 months of age), and the time point at which analysis is done.

Rates and rate differences will be calculated for mortality and for hospital outcomes. Incidence rates of hospital admission, based on sentinel hospitals, are influenced by access to hospital, and the availability of alternative health facilities. Rate differences are therefore inevitably context-specific and reflect the local situation. Rates and rate differences for hospital outcomes refer to the rates of

admission to the sentinel hospitals from the defined source populations. Rate differences therefore represent a minimum bound on the rate difference that would be observed if there was uniform access to hospital and there were no competing hospitals.

Rate ratios

In practice for rate ratios, the person time will cancel out. It is only the ratio of person time in group A and group B that is relevant (rather than the total person time for the cluster) and as we are unlikely to have cluster-specific data about this ratio, we assume it is the same in each cluster within a country.

After cancelling the person time, the ratio of the total events in group A to group B is $R_1=n_{1,A}/n_{1,B}$ in the RTS,S area and $R_0= n_{0,A}/n_{0,B}$ in the comparator area.

(We could have calculated the ratios within each cluster and then compared the mean or geometric mean between the two arms, this would give equal weight to each cluster (and there would need to be an adjustment to allow for clusters with zero events in group B). The approach we will use for the MVPE weights clusters according to their population).

The variance of R_j is [23,22]:

$$V(R_j) = \left(\frac{m_j}{(m_j-1)n_{j,B}^2}\right)\sum_{i=1}^{m_j}\left(n_{j,i,A} - R_j n_{j,i,B}\right)^2 \qquad j = 0,1$$

where m_j is the number of clusters in arm j, R_j is the point estimate of the rate ratio in area j, and $n_{j,i,A}$ and is the number of events in group A in cluster i in area j and $n_{j,i,B}$ is the number of events in group B in cluster i in area j, and $n_{j,B}$ the total events in group B in area j.

The log rate ratio comparing the RTS,S and comparator areas is

$D = \log(R_1)-\log(R_0)$, with variance $V(D) = V(R_1)/R_1^2 + V(R_0)/R_0^2$.

We will have an estimate of D for each country, D_1, D_2 and D_3, the combined estimate is then $D_{combined} = \sum D_i/V(D_i)/\sum 1/V(D_i)$, i=1..3, and the variance is $V(D_{combined})= 1/\sum[1/V(D_i)]$.

To test for interaction by country (Cochran's Q test), $\sum[(D_{combined} - D_i)/V(D_i)]^2$ is referred to the X^2 distribution with 2 degrees of freedom [24].

The final rate ratio is given by $\exp(D_{combined})$ and the 100(1-α)% confidence interval by

$\exp[D_{combined} +/- t_{\alpha/2,C-6}\sqrt{V(D_{combined})}]$, with df equal to the total number of clusters C less 2x3=6.

This final stratified rate ratio represents *the average increase or decrease in incidence of the outcome due to introduction of RTS,S vaccine, in the age group of children eligible to receive the vaccine, across all three countries*. As for the rate differences, the estimated value will depend on the analysis population chosen (A1, eligible for RTS,S; A2 eligible for 3 doses of RTS,S; or A3 eligible for 3 doses and followed up to 39 months of age) and the time point when analysis is done.

It is not planned to adjust for cluster-level covariates. (The baseline surveys are not designed to produce accurate estimates of malaria transmission and other indicators for individual clusters). We will adjust analyses of impact for transmission strata, in a secondary analysis.

Gender interaction

The approach in the preceding section can be used to compare the effect of RTS,S on mortality rate in girls and boys. If group A is redefined to mean girls eligible to receive RTS,S, and group B redefined to mean boys eligible to receive RTS,S vaccine, then R_1 in the table is the female:male mortality ratio in the RTS,S area and R_0 the female:male mortality ratio in the comparator area. Assuming the ratio of person time (girls:boys) is the same in both areas of the evaluation, then exp($D_{combined}$) represents the ratio of the female:male mortality ratio in the RTS,S area to the female:male mortality ratio in the comparator area. However this ignores the data in group B in each cluster. Group B data can be included to adjust for any randomization imbalance in the female:male mortality ratio:

$g_{j,i,A}$ is the number of deaths among girls in group A in cluster *i* in area *j* and $b_{j,i,A}$ is the corresponding number of deaths among boys. The total events in girls in group A in area *j* is $g_{j,A}$ and in boys $b_{j,A.}$ The corresponding number in group B are $g_{j,i,B}$ and $b_{j,i,B}$ and the totals $g_{j,B}$ and $b_{j,B.}$ The female:male mortality ratio in group A is $R_{j,A}= g_{j,A}/b_{j,A}$ and in group B, $R_{j,B}=g_{j,B}/b_{j,B}$. The double ratio ($R_{j,A}/R_{j,B}$) has variance [23, p. 183-4]:

$V(R_{j,A}/R_{j,B})=(R_{j,A}/R_{j,B})^2 \{ V(R_{j,A})/R_{j,A}^2 + V(R_{j,B})/R_{j,B}^2 - 2.Cov(R_{j,A}.R_{j,B})/(R_{j,A}.R_{j,B}) \}$, j=0,1

where $V(R_{jk}) = \left(\frac{m_j}{(m_j-1)b_{j,k}^2}\right)\sum_{i=1}^{m_j}(g_{j,i,k} - R_{j,k}b_{j,i,k})^2 \quad j = 0,1; k = A,B$

and $\text{cov}(R_{j,A}R_{j,B}) =$

$$\frac{m_j}{(m_j-1)b_{j,A}b_{j,B}}\sum_{i=1}^{m_j}(g_{j,i,A}g_{j,i,B} - R_{j,A}g_{j,i,B}b_{j,i,A} - R_{j,B}g_{j,i,A}b_{j,i,B} + R_{j,A}R_{j,B}b_{j,i,A}b_{j,i,B})$$

The log of the ratio comparing the female:male mortality ratios in RTS,S and comparator areas is

$W=\log(R_{1,A}/R_{1,B}) - \log(R_{0,A}/R_{0,B})$, with variance $V(W)= V(R_{1,A}/R_{1,B})/(R_{1,A}/R_{1,B})^2 + V(R_{0,A}/R_{0,B})/(R_{0,A}/R_{0,B})^2$,

and the 95% confidence interval for the ratio is exp(W +/- $t_{\alpha/2,C-2}\sqrt{V(W)}$). For pooled analysis over the three countries, as before we will have an estimate of W for each country, W_1, W_2 and W_3, the combined estimate is then $W_{combined} = \sum W_i/V(W_i)/\sum 1/V(W_i)$, i=1..3, and the variance is $V(W_{combined})= 1/\sum[1/V(W_i)]$. The final ratio of female:male mortality ratios is given by exp($W_{combined}$) and the 100(1-α)% confidence interval by exp[$W_{combined}$ +/- $t_{\alpha/2,C-6}\sqrt{V(W_{combined})}$], with df equal to the total number of clusters C less 2x3=6.

As the data of the phase 3 trial suggested stronger effects in older children, the comparison will be done for all vaccine-eligible age groups, and separately for older eligible age-groups.

These analyses will exclude deaths due to injury.

21. Databases

In each country, ODK, Redcap or a mixture of both are being used to collect and manage evaluation data. The tools used are tested to ensure integrity of the data – including links between levels of data and appropriate range checks. The master databases are held on secure servers in each of the institutions. The database servers used are MySQL and are linked into the ODK/Redcap systems by the applications themselves.

Each institution then has a separate database (for each study/survey) that is populated from the corresponding master database. All cleaning takes place in this second database. Each country has developed queries that are run on the collected data to check for completeness, duplicates and consistency.

These data are intended to be flagged to indicate the record's completeness – so that data still being reviewed can be excluded from any on-going analyses (especially relating to safety reporting.

The data (complete data, or a subset including only variables specified in the protocol) are delivered to WHO each month (on the 15th, complete to the end of the preceding month). WHO then review data delivered to look at consistency and completeness relating to key outcomes of safety and impact. These reviews are fed back to country teams to help them target improvements in data capture.

EPI data: In Malawi, monthly data for each health facility, on routine vaccine delivery, extracted from the DVDMT system. They are migrating to DHIS2 and the facility level data will then be extracted from DHIS2 by the MVIP data manager. In Kenya, DHIS2 is being used for EPI data, RTS,S variables are being incorporated. Facility level data will be extracted by the MVIP data manager. In Ghana, DHIS2 is being used, access to these data is being requested.

AEFI and AESI: Mechanisms for access to data on AEFI and AESI in each country are being developed.

22. Data sharing plans

Although plans are not finalised it is envisaged that data sets be lodged with an appropriate data archiving service. We will ensure that the data lodged in the repository will be anonymised, fully documented, and in a format that can be reused by the widest audience, and in compliance with requirements of journals in which results are to be published.

The choice of public repository should be determined by WHO policy.

The terms under which data are shared will be developed using best practice of the WHO's archiving services.

Data sharing agreements, setting out the terms under which research data are to be shared, will be developed.

23. Costs and cost effectiveness

A costing analysis will be conducted as a sub-study within the broader evaluation of the feasibility of RTS,S/AS01 introduction. The costing study will evaluate the cost of introducing and delivering RTS,S/AS01 in each of the pilot countries (Ghana, Kenya, and Malawi) through the analysis of the cost of introduction (start-up cost) and delivery (recurrent cost) of RTS,S/AS01 and the generation of evidence on the cost implications of RTS,S/AS01 introduction and delivery. Understanding these costs will be critical in assessing economic feasibility of delivering RTS,S/AS01 as well as help inform decision making and planning around further use of the vaccine.

Specific objective for costing analysis: The costing study specifically aims to estimate the incremental cost of introducing and delivering RTS,S within the MVIP pilot areas through the routine immunization programs.

Approach*:* The analysis of costs will be done from the governments' perspective and will only include the incremental cost of introducing and delivering the RTS,S vaccine into the existing immunization programs in each of the three countries. The study will take an activity-based costing approach where all activities associated with the introduction and delivery of vaccine are identified and costed individually. Both financial and economic costs associated with be included in the analysis.

The costing analysis will gather and utilize actual data on activities, costs/expenditure, and coverage (administrative) data from the pilot introduction in three countries. Data on activities, costs, and outcomes will be gathered from health administrative units and facilities at all levels (national and sub-national) of the countries.

- At the National level, data collection will occur at the immunization program (EPI) and the central level vaccine store. Data on vaccine storage and supply chain costs will also be gathered from all regional level vaccine stores included in the MVIP (including Phase 4) areas.
- Data related to service delivery will be collected from a sample of health facilities. The sample frame will include all interventions clusters (districts in Ghana and Malawi; sub-counties in Kenya).
- Fifteen to twenty-five percent of clusters from the sample frame will be selected as sample clusters for data collection using stratified a random sampling method selected for each country. Stratification will be based on the following variables, as prioritized by the National level MVIP stakeholders:
    - Geography (urban/rural)
    - Population density (high/medium/low) [or expected number of doses delivered]
    - EPI performance *(e.g. high/medium/low)*
- The sampling frame will guide the selection of health facilities within each sample cluster for data collection according to facility type (for example - sub-health post, primary health care center, dispensary, if relevant), geography (urban/rural), ownership (private/public), and population served (low volume/high volume).
- Within each sample cluster, two public health facilities (ideally, two for each level type) will be purposively selected in consultation with the local health officials in country. Selection of health facilities will seek to capture variation in within cluster heterogeneity such as urban/rural differences, high/low volume facilities). At least one private sector facility will also be selected in each cluster.
- In the event country plans to deliver vaccine through community outreach and mop-up activities, data will be gathered from at least one such events in each sample sub-region.

To facilitate the costing analysis, an excel based Malaria Vaccine Introduction planning and Costing Tool (MVICT) has been developed to guide data gathering and generate cost estimates. The tool has been reviewed by WHO's Immunization and Vaccine Implementation Research Advisory Committee (IVIR-AC) and validated for each country through a preliminary analysis of the cost of continuing vaccination.

The costing analysis will only incorporate resources spent on introduction and delivery of RTS,S/AS01, excluding costs that are primarily incurred for the evaluation component of the MVPE. All MVIP related activities will be categorized as components of the immunization program, including micro-planning; training; vaccine procurement; development of information, education and

communication (IEC) and social mobilization materials; cold chain expansion; service delivery; supervision; and monitoring of vaccine delivery. Across all activities, both the value of financial and economic resources will be measured and will be further categorized as recurrent and capital costs. Table 1 shows a framework of costs of major activities for MVIP vaccination that would be categorized under recurrent or capital costs and financial or economic costs. Please note that collecting data on RTS,S/AS01 introduction activities is a part of the study data collection, and therefore the activities listed in the tables below may not be comprehensive.

Financial costs are the value of resources borne by the buyer. Given that this costing analysis takes the government's perspective in evaluating the costs, financial costs would include costs that are incurred directly by the Ministry of Health. This would include, for example, the value of actual resources purchased for the RTS,S/AS01 introduction such as injection supplies, outreach allowances and per diem, resources used in training and developing new communication materials.

Economic costs comprise the value of all outlays for the vaccine introduction as well as those already paid for by the Ministry of Health and other sources of financing, e.g. the salaries of health personnel, vaccines paid for by partners, and time of volunteers.

Recurrent costs include the value of resources incurred that last less than one year. These include operational costs of the program such as the value of personnel time, transport, maintenance, monitoring and evaluation, and supervision as well as costs of short-term training activities that last less than a year.

Capital costs include the value of resources that last longer than one year, such as cold chain equipment and vehicles. In the context of new vaccine introduction, start-up costs such as micro-planning, initial training and social mobilization/IEC material development, as well as additional cold chain equipment, vehicles and incinerators will be categorized as capital cost. Capital costs will be annualized using straight line depreciation based on the useful life of years for financial costs and annualized and discounted using a standard 3% discount rate for economic costs.

The costing analysis will follow the standard guidelines in estimating the costs RTS,S/AS01 introduction and delivery. [2], [3] For each activity category, the major cost components to be included in costing is given in Table 1. For each sub-activity, expenditure on each input resources including personnel (personnel time spent, allowances and per diems), supplies and materials (including conference packages, stationary, communication materials), and other direct costs (including transport reimbursement, fuel, venue hire, etc) will be gathered. The cost per unit for each input will be multiplied with the number of people and the number of days (participants/facilitators), as applicable, to generate cost per activity.

The key outputs of the costing analysis will be are listed below:

- Financial and economic costs per dose delivered

[2]World Health Organization. Guidelines for estimating costs of introducing new vaccines into the national immunization system. Geneva, Switzerland 2002. Available at: http://archives.who.int/vaccines-documents/DocsPDF02/www665.pdf

[3] Working Paper. Common approach for the costing and financing analyses of routine immunization and new vaccine introduction costs. NUVI). Bill and Melinda Gates Foundation. Seattle, USA 2013. Available at: https://static1.squarespace.com/static/556deb8ee4b08a534b8360e7/t/55970258e4b03cf942da51ac/1435959896232/WEBSITE_Common+Approach.pdf

- Financial and economic cost per immunized child
- Financial and economic cost per fully immunized child
- Financial and economic cost of delivery per dose delivered
- Financial and economic cost of delivery per immunized child
- Financial and economic cost of delivery per fully immunized child
- Cost drivers of vaccine introduction and delivery

The *cost per dose* will be estimated by dividing the total cost procuring, introducing, and delivering the vaccines divided by the total number of doses delivered. The *cost per fully immunized child* will be calculated by dividing the total cost by the number of children receiving all 4 doses of RTS,S. The *cost of delivery per dose* will be calculated by subtracting the procurement costs (including cost associated with shipping and handling) from the total cost and dividing it by the total number of doses delivered. For each output, both financial and economic costs will be reported.

24. Routine AEFI surveillance (Adverse Events Following Immunization)

The is the "routine" or "passive" system in all countries that uses data on all AEFI cases for all vaccines notified to the health system through spontaneous reporting from all over the country. This is collected nationwide from Health care providers, peripheral health staff, nurses, doctors etc. The primary source of data is the AEFI reporting form that is linelisted and data analyzed. Routing of data is from patient to health care provider who completes the reporting form which is then routed to the district to province and to national level. The data are maintained by the EPI Program / NRA and is shared with partners and other stakeholders.

25. AESI Surveillance (Adverse Events of Special Interest)

AESIs are a subset of adverse events following immunization that have been seen with other vaccines or have been proposed based on theoretical concerns specific to a vaccine. A harmonized list with 9 conditions have been included for all MVIP countries and the individual countries have added additional conditions to the list based on local situations. The harmonized list includes, Anaphylaxis Cellulitis, Abscess, Meningoencephalitis, Stevens-Johnson Syndrome, Major organ failure -hepatic failure -renal failure, Thrombocytopenia/ purpura, Acute Flaccid Paralysis, Allergic/Hypersensitivity reaction and Toxic Shock Syndrome.

AESI surveillance is complementing the planned Phase IV cohort studies and will monitor 60,000 subjects across three countries (10,000 subjects in vaccination areas and 10,000 subjects in non-vaccinated area). As the cohort studies will not have the power to detect rare conditions, active surveillance will be used to monitor the occurrence of AESIs. Special AESI protocol including AESI reporting formats have been developed and shared with countries to adapt locally.

The health staff from the periphery, intermediate areas and national level (as decided by the country - including AFP, measles surveillance officers FVs etc.) will do prospective surveillance & retrospective surveillance and obtain data in specific AESI forms and linelisted. Routing of data is from patient to designated staff (as decided by the approved country specific AESI guidelines) who completes the AESI reporting form which is then routed to the district, province and to national AESI database. This is maintained in a national AESI database with the EPI Program / NRA and is shared with partners and other stakeholders.

26. EPI-MAL-002 baseline study

The RTS,S/AS01 vaccine will only be implemented in malaria endemic countries of sub-Saharan Africa. Most of these countries have no baseline (i.e. before vaccine implementation) incidence data on rare diseases such as those that may be reported as Adverse Events Following Immunisation (AEFI). The EPI-MAL-002 study started in 4Q 2015 and is collecting baseline incidence data of pre-defined events, i.e. AESI, meningitis, uncomplicated, severe and cerebral malaria as well as any other condition that requires hospitalisation or leads to death. The information that will be generated through this study will be used to compare incidence rates of those events after the RTS,S/AS01 vaccine has been introduced into the EPI programme. The study is being conducted in children aged 5 years or younger living in well-defined geographical areas. Recruitment of the subjects occurs through both active (10 home visits and continuous monitoring of outpatient visits and hospitalisations at all health care facilities) and enhanced hospital surveillances (continuous monitoring of hospitalisations). The sample size targets the enrolment of about 30,000 children, of which around 20,000 will be enrolled in Ghana and Kenya study sites that will also participate in the EPI-MAL-003 study.

27. EPI-MAL-003 safety, effectiveness and impact Phase IV study

Based on an identical methodology and conducted in the same setting, the EPI-MAL-003 study is planned to monitor the onset of the same pre-defined event as in the EPI-MAL-002 study, i.e. AESI, meningitis, uncomplicated, severe and cerebral malaria as well as any other condition that requires hospitalisation or leads to death, after RTS,S/AS01 vaccine implementation by the Ministries of Health (in the framework of the MVIP). Together, data collected through the EPI-MAL-002 and EPI-MAL-003 studies will allow a temporal comparison of the occurrence of adverse and malaria events between vaccinated and unvaccinated subjects before and after vaccine implementation.

In addition, because of the WHO recommendation to introduce RTS,S/AS01 through a phased (cluster randomised) pilot implementation, the EPI-MAL-003 study will also include a concurrent comparison of the occurrence of those events between vaccinated and unvaccinated subjects living in exposed or unexposed clusters.

The EPI-MAL-003 study sample size is 22,500 children to be enrolled from clusters where RTS,S/AS01 is implemented and 22,500 children from unexposed clusters.

The analysis plans for EPI-MAL-003 are detailed in the protocol [25]. EPI-MAL-003 study will have interim results available in Sep 2023, final results April 2026, with the possibility of an intermediate analysis in mid-2021 (if this time is chosen for the MVPE intermediate analysis), including meningitis, cerebral malaria and mortality by gender.

28 References

[1] Schellenberg D (2018) An evaluation of the cluster-randomised pilot implementation of RTS,S/AS01 through routine health systems in moderate to high malaria transmission settings in Africa. Master Protocol, version 8 RTS,S_MVIP_v8 18jul18, version 9 Apr 2020. WHO: Geneva.

[2] The RTS,S Clinical Trials Partnership. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3,individually randomised, controlled trial. The Lancet (2015). doi:10.1016/S0140-6736(15)60721-8

[3] WHO (2019) Framework Malaria Vaccine Implementation Programme (MVIP) Framework for Policy Decision on RTS,S/AS01 – DRAFT v28Feb2019. PREPARED BY THE FRAMEWORK FOR POLICY DECISION ON RTS,S/AS01 WORKING GROUP AND THE WHO SECRETARIAT. Unpublished report. WHO: Geneva

[4] Randomization for the MVPE. Unpublished report.

[5] WHO (2015) Malaria vaccine background paper for SAGE. https://www.who.int/immunization/sage/meetings/2015/october/1_Final_malaria_vaccine_background_paper_v2015_09_30.pdf

[6] Cheung YB, Jeffries D, Thomson A, Milligan P (2008) A simple approach to test for interaction between intervention and an individual-level variable in community randomized trials. Tropical Medicine and International Health 13(2):247–255.

[7] Hayes RJ and Moulton LH (2017) Cluster randomized trials. 2nd edition. New York: CRC Press.

[8] Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF International. 2015. Ghana Demographic and Health Survey 2014. Rockville, Maryland, USA: GSS, GHS, and ICF International.

[9] National Statistical Office (NSO) [Malawi] and ICF. 2017. Malawi Demographic and Health Survey 2015-16. Zomba, Malawi, and Rockville, Maryland, USA. NSO and ICF.

[10] Kenya National Bureau of Statistics and ICF International (2015) 2014 Kenya Demographic and Health Survey (2014 KDHS). Nairobi, Kenya and Rockville, Maryland, USA. KNBS and ICF.

[11] United Nations (1956) Manuals on methods of estimating population. Manual III. Methods for population projections by sex and age. ST/SOA/Series A. Population Studies, No.25. United Nations Department of Economic and Social Affairs, New York, 1956.

[12] Ewbank DC (1981) Age Misreporting and Age-Selective Under Enumeration: Sources, Pattern and Consequences for Demographic Analysis. Committee on Population and Demography, Report No. 4. National Academy Press, Washington DC, 1981

[13] WHO (2015) Assessing and Improving the Accuracy of Target Population Estimates for Immunization Coverage. World Health Organization. Working Draft. Revised November 2015 https://www.who.int/immunization/monitoring_surveillance/data/Denominator_guide.pdf

[14] Alegana VA, Atkinson PM, Pezzulo C, Sorichetta A, Weiss D, Bird T, Erbach-Schoenberg E, Tatem AJ. 2015 Fine resolution mapping of population age-structures for health and development applications. J. R. Soc. Interface 12: 20150073. http://dx.doi.org/10.1098/rsif.2015.0073

[15] Webera EM, Seaman VY, Stewart RN, Bird TJ, Tatem AJ, McKeea JJ, Bhaduria BL, Moehla JJ, Reitha AE (2018) Census-independent population mapping in northern Nigeria. Remote Sensing of Environment 204 (2018) 786–798

[16] WHO (2018) Analysis and use of health facility data: Guidance for immunization programme managers. Working Document – February 2018. WHO:Geneva. https://www.who.int/healthinfo/FacilityAnalysisGuide_Immunization.pdf?ua=1

[17] WHO (2018) Vaccination Coverage Cluster Surveys: Reference Manual. Geneva: World Health Organization; 2018 (WHO/IVB/18.09). Licence: CC BY-NC-SA 3.0 IGO. https://www.who.int/immunization/documents/who_ivb_18.09/en/

[18] Hill, K., & Upchurch, D. M. (1995). Gender differences in child health: evidence from the demographic and health surveys. In Population & Development Review (1 ed., Vol. 21, pp. 127-151)

[19] United Nations, Department of Economic and Social Affairs, Population Division (2011). Sex Differentials in Childhood Mortality (United Nations publication, ST/ESA/SER.A/314). https://www.un.org/en/development/desa/population/publications/pdf/mortality/SexDifferentialsChildhoodMortality.pdf

[20] Keyfitz N and Caswell H (2005) Applied Mathematical Demography. 3rd edition. Springer: New York.

[21] Ashorn P, Maleta K, Espo M, Kulmala T (2002) Male biased mortality among 1–2 year old children in rural Malawi. Arch Dis Child 87:386–387

[22] Bennett S, Parpia T, Hayes R, Cousens S (2002) Methods for the analysis of incidence rates in cluster randomized trials. International Journal of Epidemiology 31:839-46

[23] Cochran WG (1977) Sampling techniques.3rd edition.Chichester: Wiley.

[24] Cochran WG (1954) The combination of estimates from different experiments. Biometrics 10:101-29.

[25] GSK (2017) A prospective study to evaluate the safety, effectiveness and impact of the RTS,S/AS01E vaccine in young children in sub-Saharan Africa. 115056 (EPI-MALARIA-003 VS AME).

Annex 1: Method for estimating age-specific cluster populations

Method to estimate age-specific cluster populations:

1. Calculate average number of births per month, y=annual births/12
2. Regression of y on mid-year month number (6.5, 18.5,30.5,...), using a quadratic function of month number
3. Predict the value of y for each month Jan 2018-Dec 2023
4. Rescale the monthly values within each year, so that the annual total births is correct
5. Use the mid-year estimates of the total population under 5 years of age, for the years 2018-2023, to estimate the total in each month, as follows:
6. Regression of the total under 5 on the mid-year month number (6.5, 18.5,...), using a quadratic function of month number
7. Predict the value for each month Jan 2018-Dec 2023
8. Rescale so that the mean population under 5 in each year is equal to the census estimate of the mid-year population under 5
9. Set the population aged 0 months, in each calendar month, equal to the number of births in that month, and set starting values for the population in each month of age from 1 to 59, in each calendar month, by dividing the under-5 population for that month (less the number aged 0 months) equally
10. Calculate monthly age-specific survival probabilities and apply to the populations in each diagonal in the population matrix n(a,b), =population aged a months in calendar month b:
11. n(1,1)=n(0,0)xp0, p0=1-neonatal mortality,
12. n(j,j)=n(j-1,j-1)xg, g=exp(log(f)/10), where f=(1-infant mortality)/(1-neonatal mortality), for j=2…11
13. n(j,j)=n(j-1,j-1)xk, k=exp(log(h)/48), where h=(1-under 5 mortality)/(1-infant mortality) , for j=12…59
14. To complete the lower triangle of the matrix, when row j (month of age j) is completed for all calendar months Feb 2018-Dec 2023, the first cell n(j,1) is set equal to n(j,2), before moving to the next row
15. Rescale so column totals (total population in a calendar month) are correct, while keeping age 0 unchanged
16. Repeat 11-15 until convergence

Annex 2: Example dataset (hypothetical data) and Stata code to calculate rate ratios and differences

area: evaluation area, 0-comparator 1-RTS,S; cluster: cluster number; na: number of events in group a; nb: number of events in group b; ya: person time in group a; yb: person time in group b. Column headings are the Stata variable names.

area	cluster	na	ya	nb	yb	area	cluster	na	ya	nb	yb
1	1	5	1.782	0	6.239	0	3	0	1.697	0	5.942
1	2	2	1.239	2	4.337	0	4	1	1.533	0	5.368
1	6	3	2.508	2	8.779	0	5	1	1.273	1	4.456
1	8	2	2.390	2	8.365	0	7	1	1.057	1	3.699
1	10	2	2.373	1	8.306	0	9	0	2.737	0	9.582
1	12	2	2.863	2	10.023	0	11	1	1.701	0	5.955
1	13	1	1.909	0	6.684	0	15	1	1.142	0	3.997
1	14	1	1.134	0	3.971	0	16	2	1.646	1	5.763
1	19	1	1.675	1	5.865	0	17	1	2.110	1	7.385
1	21	1	2.949	0	10.324	0	18	0	2.751	0	9.631
1	23	3	2.452	4	8.585	0	20	0	1.409	0	4.932
1	24	1	1.090	1	3.817	0	22	7	2.785	4	9.749
1	27	3	2.491	0	8.721	0	25	0	2.168	0	7.591
1	28	2	1.992	0	6.972	0	26	1	1.739	0	6.088
1	30	0	2.433	0	8.517	0	29	5	2.701	9	9.454
1	31	0	2.719	0	9.518						

Stata code to estimate rate ratios RTS,S: comparator, and rate differences comparator-RTS,S:

```
* rate ratio:
svyset cluster
svy:ratio control: na/nb if area==0
scalar b0=_b[control]
scalar se0=_se[control]
scalar ndf0=e(df_r)
svy:ratio rtss: na/nb if area==1
scalar b1=_b[rtss]
scalar se1=_se[rtss]
scalar ndf1=e(df_r)
scalar se=sqrt(se0*se0/(b0*b0)+se1*se1/(b1*b1))
scalar b=log(b1)-log(b0)
scalar tval95=-invt(ndf0+ndf1,.025)
scalar tval99=-invt(ndf0+ndf1,.005)
di "Rate ratio:" exp(b),"95%CI:" exp(b-tval95*se),exp(b+tval95*se), ///
     "99%CI:" exp(b-tval99*se),exp(b+tval99*se)
* test of interaction:
scalar t=abs(b/se)
scalar p=ttail(ndf0+ndf1,t)*2
di t,ndf0+ndf1,p
* rate difference:
reshape long n y,i(cluster) j(gp) str
encode gp,gen(ngp)
svy:ratio n/y if area==0,over(ngp)
lincom _b[a]-_b[b]
scalar d0=r(estimate)
scalar sed0=r(se)
scalar ndf0=r(df)
svy:ratio n/y if area==1,over(ngp)
lincom _b[a]-_b[b]
scalar d1=r(estimate)
scalar sed1=r(se)
scalar ndf1=r(df)
scalar rd=d0-d1
scalar serd=sqrt(sed0^2+sed1^2)
di "rate difference:" rd,"95%CI:" rd-tval95*serd,rd+tval95*serd, "99%CI:" ///
     rd-tval99*serd,rd+tval99*serd
```

Using the example data, this code gives the following estimates:

```
Rate ratio: 1.5650794 95%CI:0.61358341,3.992079
                      99%CI:0.44306865,5.5284287
Interaction: t=0.97838205, 29df, P=0.3359772
Rate difference: -0.15947362 95%CI:-0.66267583,0.34372859
                             99%CI:-0.8376466,0.51869936
```

Annex 3: Summary of outcomes and analysis methods

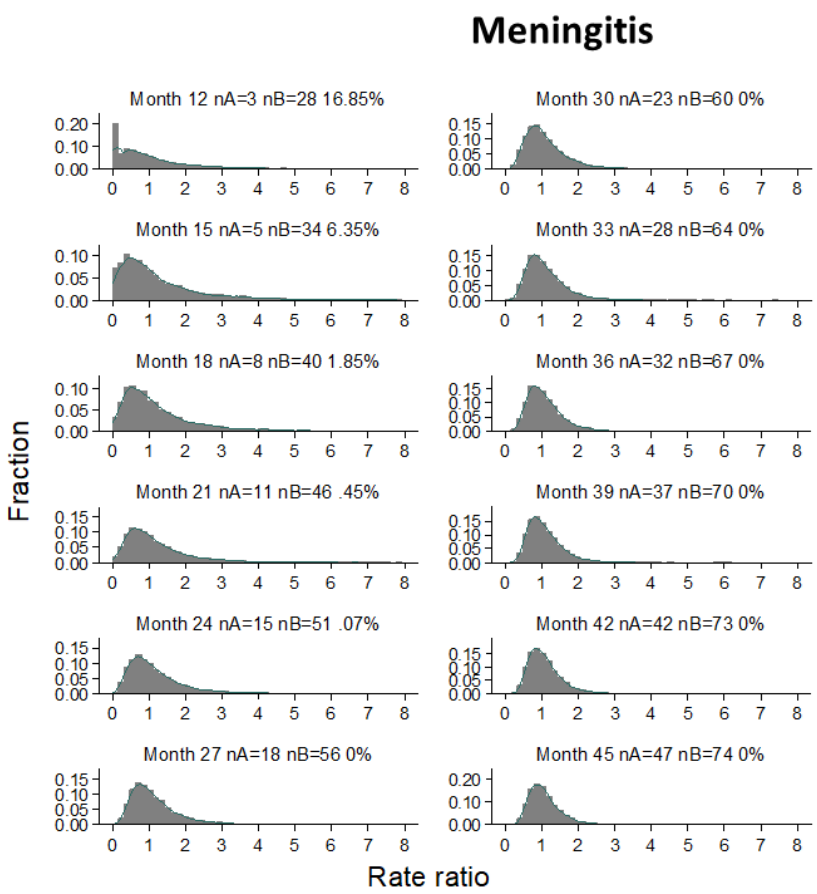
Objectives	Outcomes	Methods	Timing	Analysis population
Impact: effect of RTS,S introduction on:				
All-cause mortality	Mortality (excluding injury)	Rate ratio and rate difference	When intermediate analysis is done (driven by accrual of safety outcomes), and at 46 months	A2
Severe malaria	Hospital admission with severe malaria	Rate ratio and rate difference (cases in hospital catchments)		A2
Secondary outcomes	Severe malaria anaemia; transfusions; all-cause admissions; malaria admissions; non-malaria admissions	Rate ratio and rate difference (cases in hospital catchments)		A2
Safety: effect of RTS,S introduction on rate of admission with:				
Meningitis	Hospital admission with probable meningitis	Rate ratio and rate difference (cases in hospital catchments)	When 80-100 cases have accrued in total in the 1-59 age group	A1
Cerebral malaria	Hospital admission with cerebral malaria	Rate ratio and rate difference (cases in hospital catchments)	to be defined	A1
Evidence of differential effect on mortality by gender	Mortality (all causes excluding injury)	Interaction by gender (comparison of proportion	When about 2000-2500 deaths (excluding injury) have accrued, among vaccine-eligible age groups	A1

		boys and girls vaccinated, among deaths in vaccine-eligible age groups in RTS,S clusters)	

Objectives	**Outcomes**	**Methods**	**Timing**
Feasibility:			
Uptake of 3 doses by 12 mnths of age	Coverage of RTS,S doses 1,2,3	Ratio estimator (survey weighted); Kaplan Meier	When midline survey has been completed
Uptake of 4 doses by 27 mnths of age	Coverage of RTS,S doses 1,2,3,4	Ratio estimator (survey weighted); Kaplan Meier	When endline survey has been completed
Impact of RTS,S on EPI uptake	Coverage of penta-3 etc	Ratio estimator (survey weighted); Kaplan Meier	When midline survey has been completed
EPI data:			
Coverage of dose 4	Number of doses administered, by antigen and dose number	Descriptive	When intermediate analyses are done, and at month 46
Impact of RTS,S on EPI delivery	Number of doses administered, by antigen and dose number	Comparison of rates and drop-off rates between areas	

Annex 4: Estimates of the power of the evaluation, from simulations

A4.1 Predicted range of point estimate for the meningitis incidence rate ratio between RTS,S/AS01 and comparator areas, assuming no effect of RTS,S/AS01 on meningitis.



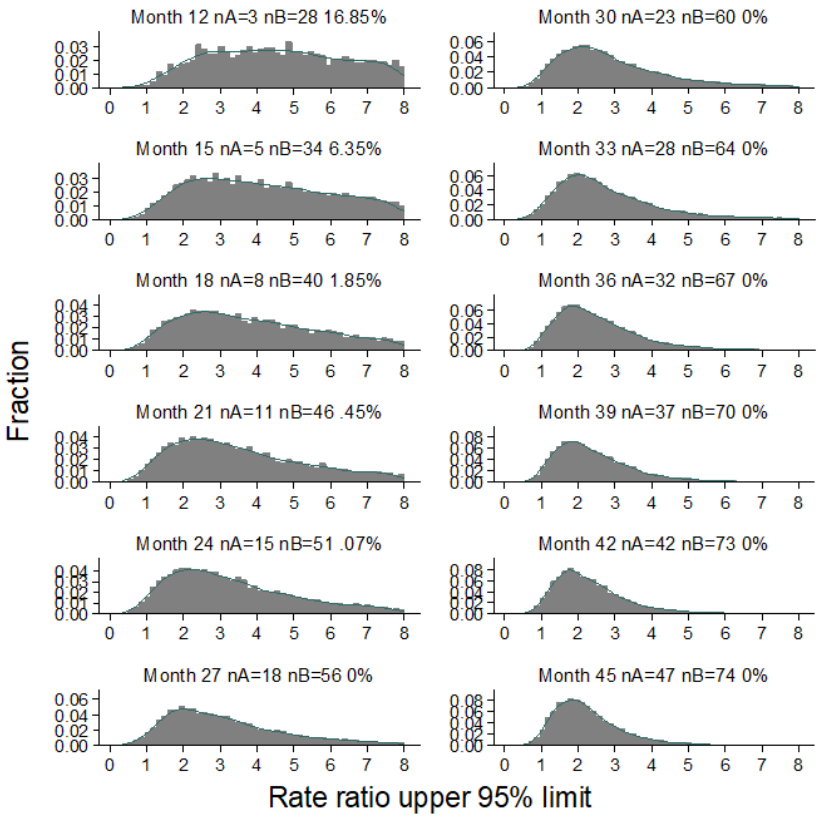
Predictions of the range of point estimates of the meningitis rate ratio, if

- the rate of admission with meningitis is 0.02/1000,
- assuming no effect of RTS,S/AS01
- data for 3 countries combined,

by months after start in first country

nA - no cases eligible
nB - no cases non-eligible
% - prob. of no cases in group B hence not estimable

A4.2 Predicted range of the upper 95% confidence limit for the meningitis incidence rate ratio between RTS,S/AS01 and comparator areas, assuming no effect of RTS,S/AS01 on meningitis.



Predictions of the range of the upper 95% confidence limit for the meningitis rate ratio, if

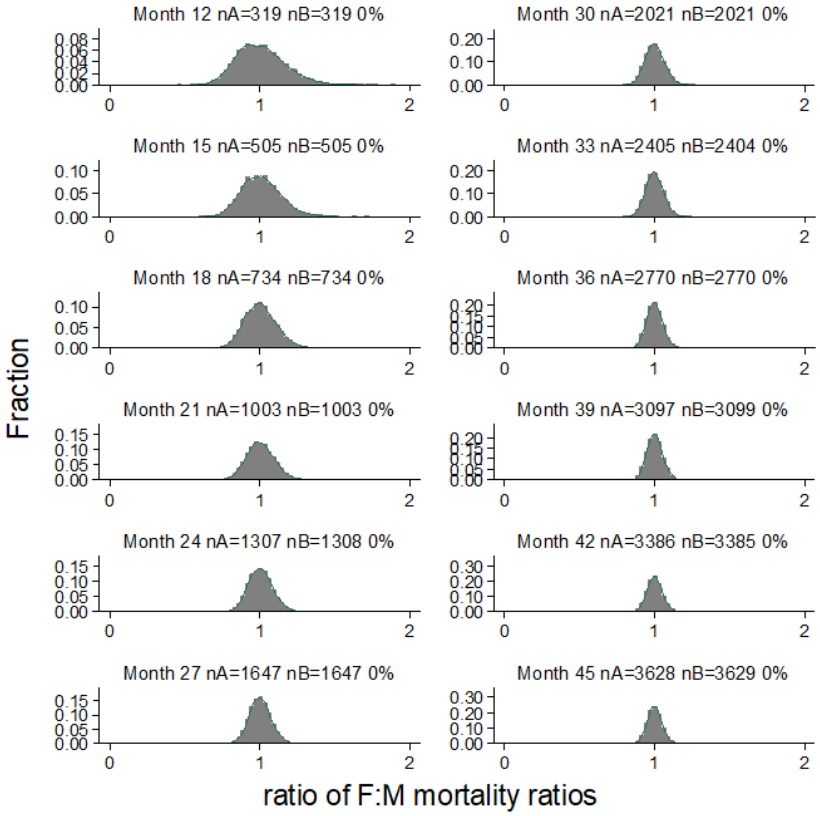
- the rate of admission with meningitis is 0.02/1000,
- assuming no effect of RTS,S/AS01
- data for 3 countries combined,

by months after start in first country

Smallest rate ratio the study would allow us to exclude:

	Power	
Month	**80%**	**90%**
12	24.7	34.6
15	17.0	29.9
18	10.5	18.3
21	7.4	11.2
24	5.8	8.2
27	4.8	6.4
30	4.2	5.4
33	3.7	4.7
36	3.4	4.2
39	3.2	3.9
42	3.0	3.7
45	2.9	3.5

A4.3 Predicted range of point estimates for the ratio of female:male mortality ratios between RTS,S/AS01 and comparator areas, assuming no differential effect of RTS,S/AS01 on survival of boys and girls.



Comparison of female:male mortality ratio between RTSS/AS01 and comparator areas:

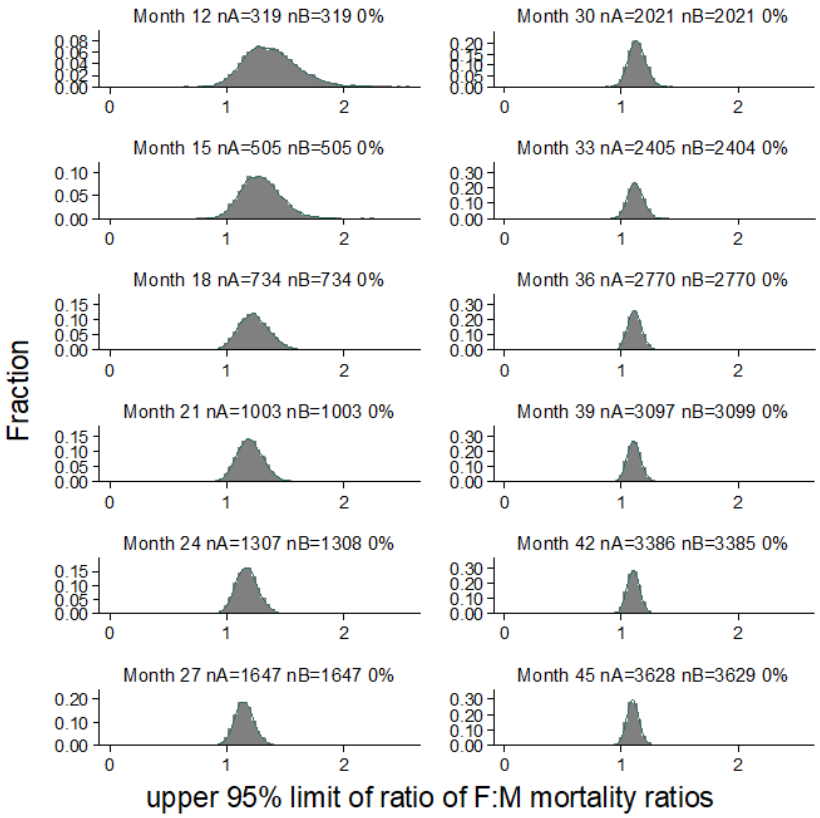
Predictions of the range of point estimates of the ratio of mortality ratios, assuming:

- mortality 2/1000 (5-23 mths), decreasing with age
- no differential effect of RTS,S/AS01
- data for 3 countries combined

by months after start in first country

nA – no. deaths girls
nB – no. deaths boys
among vaccine-eligible age groups

A4.4 Predicted range of the upper 95% confidence limit for the ratio of female:male mortality ratios between RTS,S/AS01 and comparator areas, assuming no differential effect of RTS,S/AS01 on survival of boys and girls.



Comparison of female:male mortality ratio between RTSS/AS01 and comparator areas:

Predictions of the range of the upper 95% confidence limit for the ratio of mortality ratios

nA – no. deaths girls
nB – no. deaths boys
among vaccine-eligible age groups

For meningitis, the10-fold increase in risk in vaccinated children observed in the phase 3 trial, would correspond to a rate ratio of about 5 to 6 after allowing for expected levels of coverage and contamination. The minimum number of cases required for 90% power to detect an increase in incidence due to RTS,S introduction, if the true rate ratio is 5, is about 80 to 100 cases in the vaccine-eligible and non-eligible age groups combined, Figure A4.5. (The number of events needed, is not the same at each time point, because the power depends on the relative number of cases in vaccine-eligible and non-eligible age groups and is optimal when these are equal).

Figure A4.5. Number of meningitis cases required for 90% power to detect a 5-fold increase in incidence due to RTS,S introduction, for analyses at different timepoints.

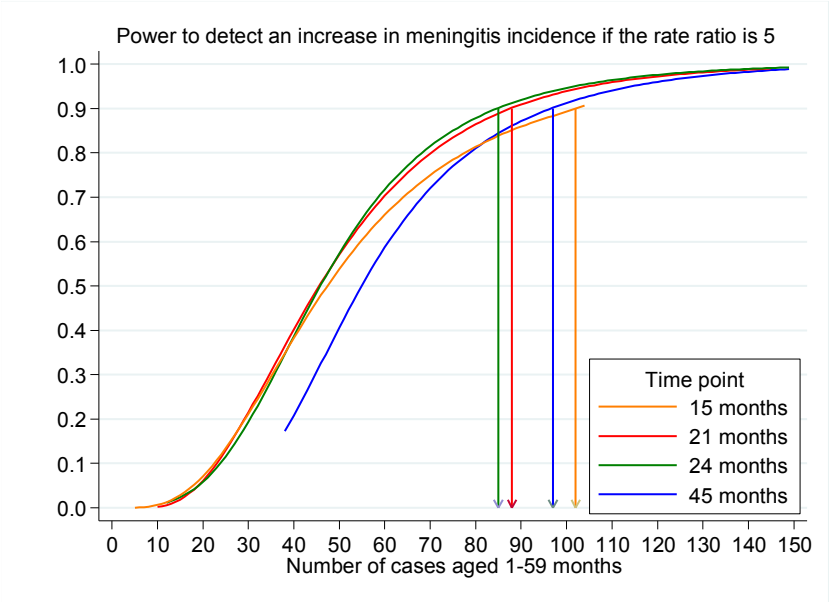
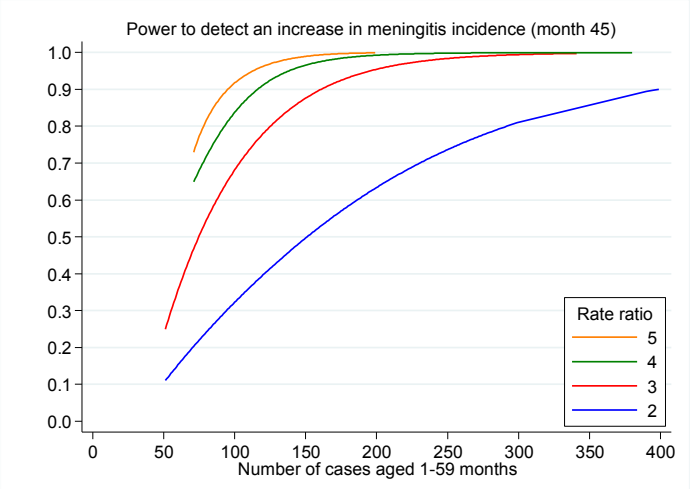


Figure A4.6. Number of meningitis or cerebral malaria cases required for 90% power to detect a 2- to 5-fold increase in incidence due to RTS,S introduction (for an analysis at the end of the evaluation).



In the phase 3 trial, among girls, mortality was 2-fold higher in those who received RTSS/AS01 than in those who received the control vaccine (risk ratio 2.0), while among boys, mortality was lower in those who received RTSS/AS01 (risk ratio 0.76), giving a ratio of mortality ratios of 2.62. This gender difference appeared stronger in the latter part of the trial, when, among girls, mortality was 3.4-times higher in RTSS recipients than controls, while in boys mortality was slightly lower (risk ratio 0.88), giving a ratio of mortality ratios of 3.84.

Mortality ratios in the phase 3 trial

	Whole period	0-20mths	21+mths
Mortality risk ratio girls:boys, RTS,S group	1.50	1.75	1.12
Mortality risk ratio girls:boys, control group	0.57	0.72	0.29
Ratio RTS,S:control	2.62	2.43	3.84
Ratio of mortality risk RTS,S:control, in girls	2.00	1.70	3.40
Ratio of mortality risk RTS,S:control, in boys	0.76	0.70	0.88
Ratio girls:boys	2.62	2.43	3.84

(These risk ratios are approximately equal to mortality rate ratios)

If RTS,S/AS01 coverage in the MVPE is 60% in both boys and girls, a ratio of mortality ratios of 2.6 translates to a ratio, at population level, of 1.9. A ratio of 2.0 among RTSS recipients would translate to a ratio of 1.6, and a ratio of 1.5 to a ratio of 1.3.

In the control group of the trial there was lower mortality in girls than boys, especially in the latter part of the follow-up. The female:male mortality ratio was 0.72 in the first 20 months of follow up (when the youngest in the cohorts were followed from age 5 months to 25 months, and the oldest from age 17 months to age 37 months). The female:male mortality ratio was 0.29 in the subsequent period, when the cohorts were aged 2 to 4 years. We can compare these female:male mortality ratios with those estimated in the UN (2011) study [19], from a variety of sources, for sub-Saharan Africa in the first decade of the 2000s: 0.85 in infants and 0.93 in children 1-4 years. Thus in the phase 3 trial, in the control group, mortality was surprisingly low in girls relative to boys.

In the RTS,S/AS01 group of the trial, mortality was higher in girls than boys, risk ratio girls:boys of 1.50 (95%CI 1.03,2.19) overall, and 1.75 (95%CI 1.09,2.80) in the earlier part of the trial and 1.12 (95%CI 0.60,2.12) in the latter part. If we view the rates in the control group of the trial as being atypical, and consider only the deaths in the RTSS group, there is a 1.5-fold excess in girls, it could be argued that we should consider the power for detecting or excluding a 1.5-fold effect.

The number of deaths in vaccine-eligible age groups required for given levels of power to detect population-level effects between 1.3-fold to 1.6-fold, is shown in Figure A4.7.

Figure A4.7: Power to detect a relative difference in female:male mortality ratio in relation to number of events accrued.