

Framework for Evaluating the Health Impact of the Scale-Up of Malaria Control Interventions on All-Cause Child Mortality in Sub-Saharan Africa

Yazoume Yé,^{1*} Thomas P. Eisele,² Erin Eckert,³ Eline Korenromp,^{4,5} Jui A. Shah,¹ Christine L. Hershey,³ Elizabeth Ivanovich,⁶ Holly Newby,⁷ Liliانا Carvajal-Velez,⁸ Michael Lynch,⁹ Ryuichi Komatsu,¹⁰ Richard E. Cibulskis,¹¹ Zhuzhi Moore,¹² and Achuyt Bhattarai⁹

¹MEASURE Evaluation, ICF, Rockville, Maryland; ²Center for Applied Malaria Research and Evaluation, Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana; ³President's Malaria Initiative, Bureau for Global Health, United States Agency for International Development, Washington, District of Columbia; ⁴Avenir Health, Geneva, Switzerland; ⁵Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, The Netherlands; ⁶United Nations Foundation, Washington, District of Columbia; ⁷Independent Consultant based on Stockholm, Sweden; ⁸Division of Data, Research, and Policy, Data and Analytics Section, United Nations Children's Fund, New York, New York; ⁹President's Malaria Initiative, Malaria Branch, Division of Parasitic Diseases and Malaria, Centers for Disease Control and Prevention, Atlanta, Georgia; ¹⁰The Global Fund to Fight AIDS, Tuberculosis and Malaria, Geneva, Switzerland; ¹¹Global Malaria Programme, World Health Organization, Geneva, Switzerland; ¹²IBTCI, Vienna, Virginia;

Abstract. Concerted efforts from national and international partners have scaled up malaria control interventions, including insecticide-treated nets, indoor residual spraying, diagnostics, prompt and effective treatment of malaria cases, and intermittent preventive treatment during pregnancy in sub-Saharan Africa (SSA). This scale-up warrants an assessment of its health impact to guide future efforts and investments; however, measuring malaria-specific mortality and the overall impact of malaria control interventions remains challenging. In 2007, Roll Back Malaria's Monitoring and Evaluation Reference Group proposed a theoretical framework for evaluating the impact of full-coverage malaria control interventions on morbidity and mortality in high-burden SSA countries. Recently, several evaluations have contributed new ideas and lessons to strengthen this plausibility design. This paper harnesses that new evaluation experience to expand the framework, with additional features, such as stratification, to examine subgroups most likely to experience improvement if control programs are working; the use of a national platform framework; and analysis of complete birth histories from national household surveys. The refined framework has shown that, despite persisting data challenges, combining multiple sources of data, considering potential contributions from both fundamental and proximate contextual factors, and conducting subnational analyses allows identification of the plausible contributions of malaria control interventions on malaria morbidity and mortality.

INTRODUCTION

Malaria remains a public health challenge in sub-Saharan Africa (SSA) despite concerted efforts from national and international partners to scale up effective malaria control interventions, such as insecticide-treated nets (ITNs), indoor residual spraying (IRS), diagnostics, prompt and effective treatment of malaria cases, and intermittent preventive treatment during pregnancy (IPTp). The World Health Organization (WHO) estimates that between 2001 and 2013, 670 million fewer cases and 4.3 million fewer malaria deaths occurred globally than would have occurred if incidence and mortality rates had remained unchanged since 2000. Of the estimated 4.3 million deaths averted between 2001 and 2013, 3.9 million, or 92%, were among children under 5 years of age in SSA.¹

Global investments in malaria control and elimination efforts exceeded USD \$2.7 billion in 2013,¹ resulting in a strong desire from funding partners and national malaria control programs for rigorous evaluations of scaling up malaria control interventions to quantify the impact on malaria-associated morbidity and mortality. Several challenges complicate this task. First, nearly all malaria control interventions are either implemented as national coverage programs or targeted to high-transmission risk areas.^{2,3} In such cases, a contemporaneous control group, or counterfactual, is often unavailable, hindering direct causal inference between exposure to malaria

control interventions and any observed changes in malaria health outcomes.⁴ Second, data on malaria morbidity, mortality, and transmission are often unavailable at the required scale or are inappropriate for use in analysis. Third, other factors can affect malaria health outcomes, such as additional maternal and child health interventions, socioeconomic conditions, and environmental factors, but they may not be well documented. Fourth, it is difficult to define an evaluation's baseline, midline, and endline,^{4,5} so that intervention scale-up timing and coverage level obtained must be carefully considered to ensure a complete assessment. Finally, the efficacy of malaria interventions is already proven, unlike other development fields, the evaluation question is more about program implementation of proven interventions and the effectiveness.

A rigorous and valid evaluation design must address these constraints; therefore, in 2007 Roll Back Malaria's (RBM) Monitoring and Evaluation Reference Group (MERG) proposed a framework for evaluating full-coverage malaria control intervention impacts on morbidity and mortality in high-burden SSA countries,⁶ which was the first attempt to propose an evaluation method that fits malaria-control intervention. That framework, however, was theoretical and not yet implemented to uncover potential challenges.

Because standard academic definitions of *evaluation* attribute change in impact measures directly to program interventions, the framework is not, in the strictest sense, an impact evaluation. It is more about plausibility design as opposed to adequacy and probability design.⁷ In this article, impact evaluation refers to the potential contribution of a package of malaria control interventions, which could be

* Address correspondence to Yazoume Ye, MEASURE Evaluation, ICF, 530 Gaither Road, Suite 500 Rockville, MD 20850. E-mail: yazoume.ye@icf.com

TABLE 1
Definition of key terms used in this paper

Terms	Definition
All-cause child mortality rate	Probability of dying from any cause between the first and fifth birthday per 1,000 children who survived to age 12 months ¹⁷
Civil registration and vital statistics	A system for recording vital events in a population, including births and deaths, with medical certification of the cause of death according to the rules and procedures of the International Classification of Diseases
Confirmed malaria case	Suspected malaria case in which malaria parasites have been demonstrated in a patient's blood by microscopy or a rapid diagnostic test ³²
Contextual factors	Non-malaria programs and other factors, such as rainfall, socioeconomic status, urbanization, and policy changes, that could confound the association between scale-up of the intervention and its potential health impact or modify the effect of the intervention, and affect the conclusion
Impact evaluation	Within the context of this paper, impact evaluation refers to the potential contribution of a package of malaria control interventions, which could be considered the national malaria program, to a given outcome
Malaria parasitemia	Presence of malaria parasites in the blood or number of parasites per volume of blood
Malaria parasite prevalence	Proportion of children ages 6–59 months with malaria parasite infection ¹⁷
Malaria transmission	Spread of malaria by completion of a full transmission cycle (man→mosquito→human)
Malaria transmission intensity (force of infection)	Measured as entomological inoculation rate (EIR): the number of infectious mosquito bites a person is exposed to in a certain time period, typically a year
Malaria-specific mortality	Deaths in which malaria was the underlying cause. The World Health Organization (1993) defines it as “the disease or injury which initiated the train of morbid events leading directly to death”
Plausibility argument	An assumption that mortality reductions can be attributed to programs if improvements are found along the causal pathway between intervention scale-up and mortality trends ⁷
Population-level malaria morbidity indicators	Indicators on malaria morbidity collected through population-based surveys; examples are malaria parasite prevalence and anemia
Under-five mortality	Probability of dying before the fifth birthday per 1,000 live births
Verbal autopsy	A method for determining cause of death. A knowledgeable person in the household where a deceased person lived is asked about signs and symptoms of the terminal illness, usually 1–6 months after the death. ^{18,22,28} To attribute causes of deaths, interviews are analyzed by an algorithm or clinicians who decide on causes by majority vote ⁶

national malaria programs, to change malaria morbidity and mortality (Table 1, definition of key terms).

This article expands the RBM/MERG framework and highlights some of the challenges in applying this theoretical framework. In addition, we take the RBM/MERG framework further by including secondary analyses to more rigorously examine causal inference. These analyses include stratification, use of a national-platform analytic framework, and analysis of birth history data from national household surveys to assess child survival. Our framework takes into account new program data from intervention scale-ups since 2007 and new evaluation experience.^{8,9} This updated framework may be useful to stakeholders that are responsible for implementing and evaluating malaria intervention impacts, such as ministries of health, national malaria control programs, and development partners that support malaria control efforts.

MATERIALS AND METHODS

Evaluation objectives. The main objectives of an malaria control intervention impact evaluations are to 1) measure the degree of implementation success and outcomes compared with targets set in the national strategic plan, 2) assess trends in malaria-related morbidity and all-cause under-five mortality as a complementary malaria impact indicator, and 3) assess the plausible attribution of the malaria control intervention scale-up on malaria-related morbidity and all-cause under-five

mortality, and also account for contextual factors, such as determinants of malaria transmission and child survival.

Following are the specific evaluation questions:

- Was the scale-up sufficient enough to affect national-level malaria-specific morbidity and all-cause under-five mortality?
- Was sustained coverage of malaria control intervention achieved? Was the coverage equitable across socioeconomic groups?
- Was the increase in coverage of interventions associated with impact on national all-cause under-five mortality?
- Was impact highest among those risk groups with the greatest potential to benefit from exposure to the interventions? And was the impact equitable across socioeconomic groups?
- Can other plausible explanations of the observed change in national all-cause under-five mortality be excluded?

Evaluation design. An experimental study design is assumed to be the gold standard for assessing efficacy of new interventions.⁵ Using such a design to evaluate a package of malaria control interventions is challenging for several reasons, the foremost being a lack of randomization to designate an intervention group and a control group, or a counterfactual. Where malaria control interventions are scaled up or intensified staggered over time, it may be possible to use a step-wedge quasi-experimental evaluation design.^{7,10,11} However, the step-wedge design is subject to variations among wedges

regarding geographical settings, population, or transmission patterns. The design is also subject to contamination among wedges, such as the higher burden areas may receive interventions first, compared with those with lower burden, and randomizing between the two settings poses significant ethical challenges. Also, limited subnational-level longitudinal, population-based data on the coverage of malaria control interventions and health outcomes precludes evaluation designs, such as a national platform approach, to assess a dose-response relationship between coverage and outcomes.¹¹

To address these constraints, a prepost reflexive control study design, also known as the ecological or plausibility study design, for evaluating the effectiveness of national-coverage malaria control interventions has been proposed.^{6,7,11} The plausibility study design contends that if the interventions have already proven efficacious on malaria outcomes, then it is plausible that their scale-up could have contributed to or driven the observed reductions in malaria burden, unless other contextual factors and nonmalaria related interventions are unlikely to explain all of these reductions in burden. For example, randomized controlled trials have established protective efficacy of ITN use on the incidence of malaria illness and all-cause under-five mortality;¹²⁻¹⁴ therefore, increasing coverage of ITNs use should reduce malaria morbidity and mortality over a time period consistent with the impacts established in the trials. The plausibility design is

strengthened if malaria burden baseline levels can be established before intervention scale-up, with the counterfactual being that pre-scale-up malaria trends would have continued in the absence of the interventions. Thus, over time, the intervention group serves as its own comparison group (Figure 1).

The strongest plausibility assessment can be made when there is evidence of a statistically significant increase in intervention coverage and use that is substantial enough to achieve a realistically proportional reduction in malaria-related morbidity and all-cause under-five mortality within the time frame expected (Table 2). The relationships between increasing coverage and use and reducing morbidity and mortality; however, are not linear. The plausibility assessment must also document contextual factors, both health-related and not, that may affect changes in malaria-related morbidity and child mortality.

This plausibility assessment assumes evaluators can 1) reliably measure changes in malaria intervention coverage, malaria health outcomes, and contextual factors over time; 2) determine that increases in malaria intervention coverage adequately preceded changes in malaria health outcomes; and 3) record whether health impact was highest among specific risk groups with the greatest potential to benefit from exposure to the interventions. The plausibility design should take into consideration all of the Bradford Hill criteria for causality, as outlined in Table 3.

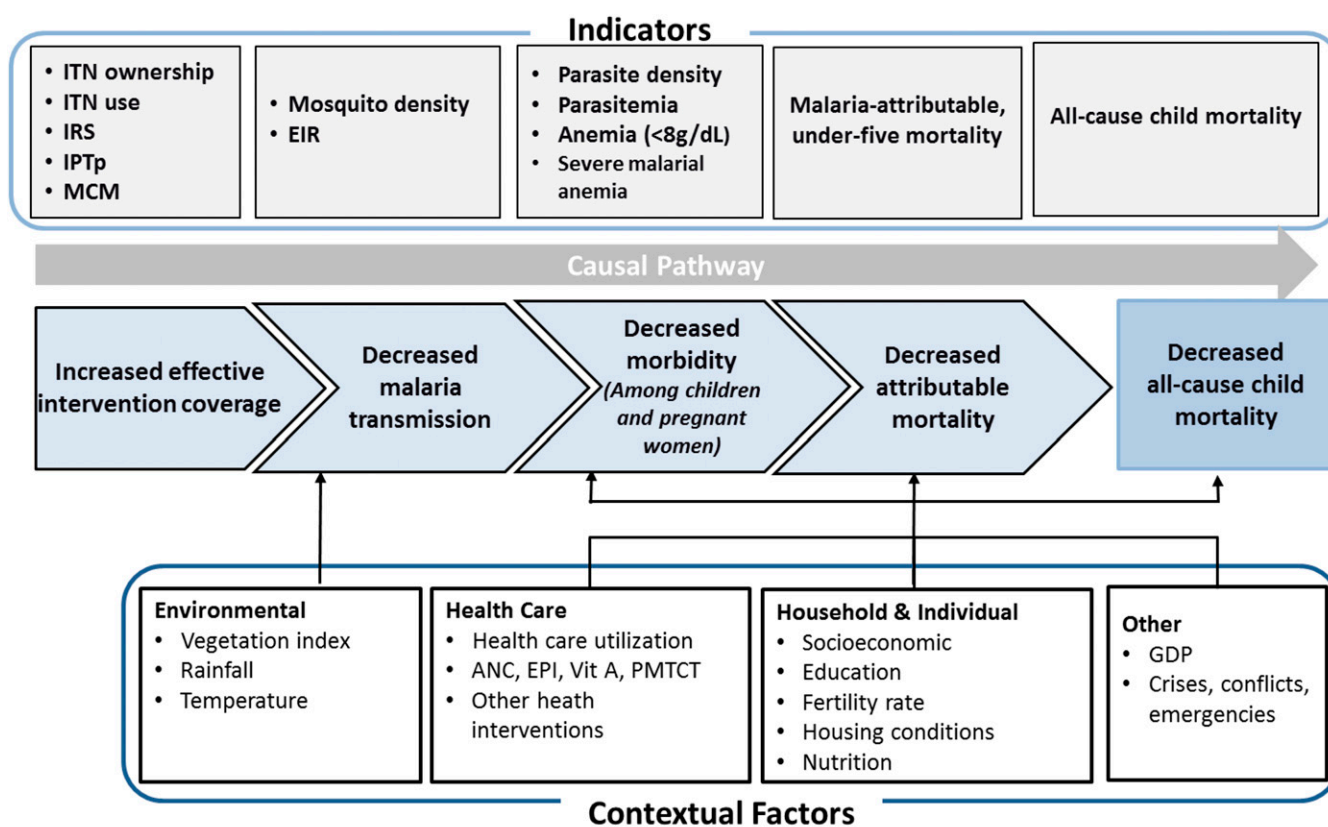


FIGURE 1. Plausibility study design framework for assessing malaria control intervention impact on malaria morbidity and all-cause child mortality. ANC = antenatal care; EIR = entomological inoculation rate; EPI = extended program for immunization; ITN = insecticide-treated net; IRS = indoor residual spraying; IPTp = intermittent preventive treatment; GDP = gross domestic product; MCM = malaria case management; Vit = vitamin; PMTCT = prevention of mother to child transmission. This figure appears in color at www.ajtmh.org.

TABLE 2
Plausibility study design strengths, limitations, and assumptions

Strengths	Limitations	Key assumptions
Intervention group serves as its own control over time	No true counterfactual, so cause and effect cannot be conclusively inferred	Program is preexisting or full (above threshold) coverage
No need to exclude any population or group from the intervention/program, so can be applied to programs with nation-wide coverage	Multiple sources of data, analyses, and triangulations needed to establish plausible impact Differential selection bias and attrition risk to cause bias and dilution of impact	Pretest (baseline) data for the relevant indicators can serve as counterfactual scenario No other plausible explanations for observed outcomes or any likely confounder effects can be adjusted for
Can adapt to use existing data collected for other purposes (DHS, MICS)	Data might not be as specific as required	All-cause under-five mortality is a sensitive, specific, and time-sensitive proxy for changes in malaria-specific mortality in highly endemic countries

Evaluation timing. Strategically timing an impact evaluation is important in achieving the most effective, informative results. Stakeholders should be enlisted to develop criteria to define the evaluation baseline, midline, and end line, based on when malaria control intervention implementation occurred and the availability of data points.⁶ When possible, baseline outcome measures should be established well before malaria control intervention scale-up because under-five mortality estimates from household surveys reflect the situation an average of 2.5 years preceding the survey. Figure 2 illustrates a sample evaluation timeframe, including intervention milestones and timing of data sources.

There is no simple, direct way to determine when an impact evaluation becomes appropriate—the exact time point after malaria control intervention scale-up begins or the level of

intervention coverage. If the scale-up is still in the early stages, coverage levels may still be too low, and an impact evaluation may not be recommended; however, it is difficult to state a threshold for when national-level impact is likely to become measurable. Conducting an evaluation takes substantial planning and inputs, as illustrated in the process of RBM impact evaluations led by the President's Malaria Initiative (PMI).¹⁵ A program performance review that assesses program implementation success may be suitable when impact evaluations are not.¹⁶

Evaluation indicators. *Changes in coverage of malaria control interventions.* MERG issued a set of primary population-level indicators to measure changes in coverage of key malaria control interventions.¹⁷ These indicators measure how well malaria control interventions are being

TABLE 3
Bradford Hill causality criteria, as applied to plausibility assessment

Criterion	Description	Assumptions
Strength of association	Strong associations are more likely to have causal components than weaker associations.	Associations can be measured
Consistency	Observing similar evaluation results across evaluation methods, over time, and across countries from meta-analyses increases the likelihood of causal relationships.	Results have been measure consistently over time and space
Specificity	Observing an association specific to outcomes of interest among specific groups increases the argument for causal effect.	Malaria interventions are highly likely to reduce all-cause under-five mortality, particularly among vulnerable groups
Temporality	Changes in program must precede changes in disease or coverage outcomes.	Scale-up of interventions has been measured
Gradient	Changes in disease or coverage outcomes increase the same amount for increases to program exposure or intensity.	Coverage has been measured in different geographic areas
Plausibility	Biological plausibility links exposure to intervention with health outcome.	Malaria contributes to all-cause child mortality
Coherence	Causal inference is possible only if the literature or substantive knowledge supports this conclusion	There are documented studies showing that malaria interventions affect mortality
Experiment	Causation is a valid conclusion if researchers have seen observed associations in prior experimental studies.	There are documented studies showing that malaria interventions affect mortality
Analogy	For similar programs operating, similar results can be expected to bolster the causal inference concluded.	Program context has been similar in the past

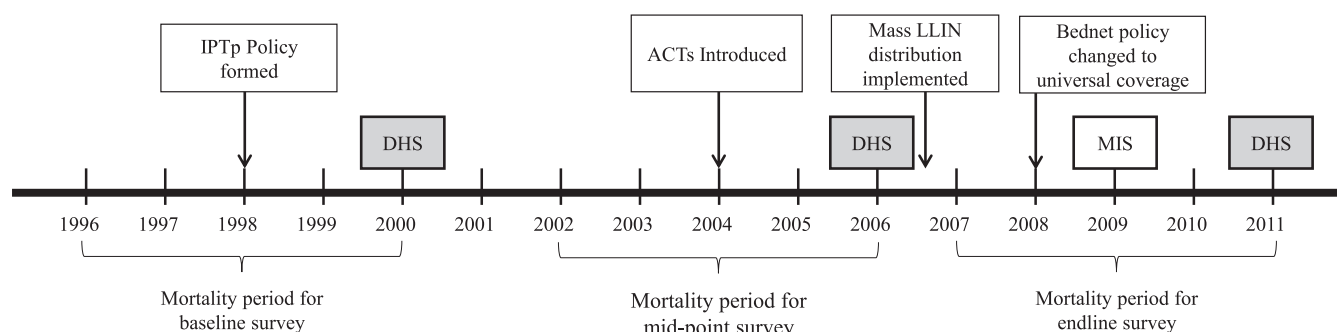


FIGURE 2. Example evaluation timeframe, from Evaluation of the Impact of Malaria Control Interventions on All-Cause Mortality in Children Under-Five in Uganda. Source: Unpublished report). DHS = Demographic and Health Surveys; MIS = Malaria Indicator Survey; ACT = artemisinin-based combination therapies; LLIN = long-lasting insecticidal nets; IPTp = intermittent preventive treatment of pregnant women.

implemented to meet coverage targets defined in national malaria control strategies.³ Key indicators include ITN ownership and use, households protected by IRS, pregnant women receiving IPTp, and children receiving effective treatment of malaria (Table 4).

Primary data sources for population-level coverage include national household surveys, such as the Demographic and Health Surveys (DHS), Malaria Indicator Surveys (MIS), and Multiple Indicator Cluster Surveys (MICS). These surveys typically are conducted every 3–5 years and provide national

TABLE 4
Key primary outcome indicators used to assess malaria control intervention scale-up

Indicator	Purpose/rationale of indicator
Vector control	
Proportion of households with at least one ITN	Measures household ITN ownership
Proportion of households with at least one ITN for every two people	Measures the proportion of households with sufficient ITNs to cover all individuals who spent the previous night in surveyed households, assuming an average of two people sharing each ITN
Proportion of population with access to an ITN in their household	Measures the proportion of the population that could have slept under an ITN, assuming each ITN is used by two people
Proportion of the population that slept under an ITN the previous night	Measures the level of ITN use among all individuals who spent the previous night in surveyed households, regardless of whether those individuals had access to an ITN in their household
Proportion of children under 5 years old who slept under an ITN the previous night	Measures the level of ITN use of children under 5 years old.
Proportion of pregnant women who slept under an ITN the previous night	Measures the level of ITN use among pregnant women
Proportion of existing ITNs used the previous night	Measures the use of existing ITNs. In certain instances, calculating the proportion of existing ITNs used the previous night will be useful for assessing the utilization of existing ITNs and determining the magnitude of nonuse of ITNs at the time of the survey
Households covered by vector control: proportion of households with at least one ITN and/or sprayed by IRS in the last 12 months	Measures the proportion of household protected by an ITN or IRS
Universal coverage of vector control: proportion of households with at least one ITN for every two people or sprayed by IRS within the last 12 months	Measures progress towards achievement of universal coverage of malaria prevention through the two main vector control activities
Intermittent preventive treatment during pregnancy (IPTp)	
Proportion of women who received three or more doses of IPTp for malaria during antenatal care visits during their last pregnancy	Measures national level coverage of use of IPTp to prevent malaria during pregnancy among women who gave birth in the last two years.
Case management	
Proportion of children under 5 years old with fever in the last 2 weeks who had a finger or heel stick	Measures national-level coverage of parasitological diagnosis among children under 5 years of age
Proportion of children under 5 years old with fever in the last 2 weeks for whom advice or treatment was sought from a formal health-care provider	Measures national-level coverage of health seeking behavior for malaria from the formal health care providers among children under 5 years
Proportion of children under 5 years old with fever in last 2 weeks who received first-line antimalarial treatment according to national policy	Measures national-level treatment coverage of children under 5 years are in accordance with national first-line malaria treatment policy.
Proportion receiving treatment with recommended first-line antimalarial, among children under 5 years old with fever in the last 2 weeks who received any antimalarial drugs	Measures what proportion of antimalarial treatment received by children under 5 years are in accordance with national first-line malaria treatment policy.

Source: Adapted from Roll Back Malaria, 2013, *Household Survey Indicators for Malaria Control*. ITN = insecticide-treated net; IRS = indoor residual spraying; IPTp = intermittent preventive treatment of pregnant women.

and subnational estimates for intervention coverage indicators. Further coverage data may come from health management information systems (HMIS), Health and Demographic Surveillance Systems,¹⁸ or special surveys conducted for specific projects, such as post-campaign long-lasting insecticidal nets coverage surveys. Other malaria program data, such as timing and quantity of malaria commodity distribution, might also help understand intervention coverage. Data from these sources, with the potential to contribute to measuring and understanding coverage, should be evaluated for rigor and incorporated in the analysis when appropriate, considering scope, sampling, and the data collection method.

Malaria morbidity, mortality, and transmission indicators. Indicators of key malaria outcomes are used to assess how the burden of malaria has changed over time, presumably as a result of malaria control intervention scale-up. Examples of these outcomes include malaria morbidity, determined by prevalence of malaria antigens or parasites; density of parasitemia (parasites/ μL) or severe anemia ($< 8 \text{ g/dL}$) in children under 5 years of age,¹⁹ or all-cause under-five mortality. When data are available, an outcome indicator can be determined from malaria transmission intensity, measured through malaria infection incidence rates, or force of infection, or the entomological inoculation rate.^{20,21}

Health Management and Information Systems have the potential to provide malaria-specific morbidity and mortality data, when proper diagnosis and medical certification of cause of death are available. The quality of the data depends on consistent, appropriate use of diagnostics and inpatient and outpatient case management. A change in diagnostic methods over the evaluation period, such as purely clinical or microscopy or rapid diagnostic test, can affect morbidity trends. In addition, health-care utilization trends and reporting rates can affect malaria morbidity and mortality calculations for health facilities. Because health-care utilization rates are low in most SSA malaria-endemic countries, only a potentially time-varying fraction of malaria deaths are likely to occur at public health facilities. Many malaria-related deaths occur at homes or at private facilities that do not link into HMIS.¹

Verbal autopsy is so far the only available and accessible method to measure malaria-specific mortality at the population level in a context where most deaths occur outside a formal health facilities. With verbal autopsy, the deceased's primary caregivers are interviewed for detailed information about events leading up to the death to ascertain and assign a cause.^{22–25} Existing verbal autopsy tools are not highly sensitive,²⁶ but efforts are underway to increase sensitivity in detecting cause of death, such as using computer-based algorithms and the simplified symptom pattern approach.²⁷ Although verbal autopsy may not provide a precise measure of malaria-specific mortality, it may give an idea of the ranking of malaria in a list of causes of death at to population level and help understand further malaria epidemiology.²⁸ In countries where such data exist, evaluations should consider exploring verbal autopsy for use in the plausibility assessment, particularly as a case study because the information in most cases is not available at the national level.

Civil registration and vital statistics systems are also potential data sources for measurement of population-level cause specific mortality.²⁹ However, in most countries data are not strong enough to use in impact evaluations at the

national level,³⁰ a circumstance that could change as confirmation of malaria diagnosis with parasite density becomes more routine and medical autopsies are performed.¹⁸

In the absence of reliable malaria-specific mortality data, and given the large indirect effects that malaria control interventions have on child health,³¹ RBM and WHO recommend using all-cause under-five mortality as the standard impact measure in malaria-endemic countries.^{17,31,32,33} This is commonly measured through standardized national household surveys, such as DHS and MICS, in most malaria-endemic countries.

Contextual factors. Contextual factors that may influence outcome or impact indicators should be investigated for their potential contribution to changes in the outcomes of interest. Where possible, these contextual factors should be directly accounted for in the analytic framework used to assess the association between changes in malaria control intervention coverage and malaria health outcomes. Contextual factors can be divided into two groups, 1) distal (e.g., gross domestic product, housing condition, maternal education, and climate) and 2) proximal (e.g., nutritional status, childhood vaccination coverage, vitamin A supplementation, and breastfeeding). Fluctuation in contextual factors is expected, so evaluators should only look for anomalies that may explain changes in all-cause under-five mortality, within the given the timeframe (Table 5).

Data analysis. The primary method to build plausibility is to assess trends in malaria intervention coverage, malaria health outcomes, and potential contextual factors. An example is the Zanzibar North A district program evaluation that incorporated data from multiple sources, including HMIS-derived monthly incidence cases as the primary outcome, and explanatory variables, such as the year of program intervention and monthly rainfall from meteorological sources, all analyzed in a negative-binomial regression framework.³⁴ In another example from Zanzibar, evaluators used log-linear regression to assess reductions in malaria cases and deaths between the preintervention period of 1999–2003 and 2008.³⁵ Evaluators concluded that significant reductions in malaria cases and malaria-related mortality were made during the period evaluated.

Another example is the evaluation of the Bioko Island Malaria Control Project³⁶ in Equatorial Guinea that assessed 4 years of high-coverage scale-up of IRS, ITNs, and improved case management on the island. Researchers used survival regression to measure the change in all-cause mortality among children under 5 years of age. Calendar year was included as a covariate to test the presence of a declining trend in child mortality. Household access to electricity was included as a proxy indicator of household wealth. A subsample was used to test the association between rainfall in the previous 12 months and child mortality. Variables of interest included possession and use of ITNs, IRS coverage, and pregnant women protected through IPTp. As hypothesized, results indicated simultaneous drops in the prevalence of malaria parasite infection, anemia, reported fever, and all-cause under-five mortality in children. This evaluation, however, did not include an in-depth discussion of contextual factors that could also have influenced the decline in all-cause under-five mortality.

The RBM impact evaluations, led by PMI, took advantage of a key strategy for data analysis: risk stratification. To

TABLE 5
Examples of contextual factors that should be examined

Category	Examples	Data sources	Justification
Child survival interventions	Expanded program on immunization coverage, such as measles and DPT3, micronutrient supplementation coverage, including vitamin A, iron, and zinc	WHO, UNICEF annual estimates of national immunization coverage UNICEF vitamin A coverage database	Observed reductions in child morbidity and mortality may actually be result of increased coverage of these programs rather than malaria control interventions.
Climatic and environmental factors	Total precipitation	DHS, MICS, MIS National meteorological agency	These factors affect mosquito breeding and malaria transmission and may cause observed changes in outcomes over time or geography, rather than the interventions themselves.
	Number of days with rain	Columbia University Earth Institute climate database	
	Land cover and vegetation	National Oceanographic and Atmospheric Administration	
	Air temperature Extreme weather events, such as floods		
Health systems factors	Per capita expenditure on health	WHO/WHOSIS	Health systems can affect comparisons across time or geography by influencing access to interventions. These factors modify the impact of malaria control interventions.
	Government expenditure on health as percentage of total government expenditure	The World Bank development indicator database	
	Availability of essential drugs Political situation and stability		
Socioeconomic factors	Household assets and income Parental education	DHS, MICS The World Bank development indicator database	If different socioeconomic groups access malaria control interventions differently, these factors may serve as effect modifier influence outcomes.
	Conflict or emergency settings GDP per capita, Gini per capita Population living below poverty line		

DPT3 = diphtheria, pertussis, tetanus vaccine, 3 doses; WHO = World Health Organization; UNICEF = United Nations Children's Fund; DHS = Demographic and Health Surveys; MICS = Multiple Indicator Cluster Surveys; MIS = Malaria Indicator Survey; GDP = gross domestic product.

strengthen the plausibility design, analyses of parasite prevalence, anemia prevalence, and mortality were stratified by age group, place of residence, and malaria risk zone, where data were available.⁸ The stratification helped identify areas and subpopulations with the greatest potential to benefit from interventions and assess whether these groups experienced the highest reductions in morbidity or mortality during the evaluation. If reduction in anemia is associated with malaria decline, we would expect to see a higher baseline and greater reduction in severe anemia prevalence in areas with relatively higher malaria risk, considering that other factors remain constant.³⁷ Furthermore, if a major part of the decline in mortality among children under five was malaria related, we would expect a greater mortality decline among children under five living in areas of greater malaria risk, compared with areas of lower malaria risk, if other factors remain constant.^{37,38} In a similar way, we would expect a greater decline in mortality among the age group most likely to experience malaria-related death (6–24 months), compared with other childhood age groups, if malaria control interventions were the main mortality reduction driver.^{39,40} However, the age groups might change with the reduction of malaria infection. In countries where significant reduction have been achieved, it might be advisable to explore older age groups.

The results from such analyses to assess trends in malaria morbidity and mortality can be strengthened using further analytic methods tailored to the available data in a specific country. Building on previous work using a national-platform approach, where districts or district-time units formed the unit

of analysis,^{41–43} an available survey data analysis in Malawi used a Poisson model at the district level to examine the association between ITN coverage and all-cause child mortality.¹² In Uganda and Senegal, evaluators conducted a regression survival analysis to examine differences in child survival before and after the scale-up began. Separately, the analysis examined trends in variables, such as mother's education, DPT3 coverage, socioeconomic growth, access to health care, and temperature and rainfall anomalies. These types of analyses can be undertaken, depending on data availability and usefulness to the objectives of the evaluation itself.

Data may also be pooled across countries to provide evidence of malaria control intervention effects at the global level (Figure 3).^{19,43} After evaluations have been completed in several countries and comparable datasets are available, this kind of meta-analysis could provide more insight into malaria program effectiveness, especially where in-country analyses lack sufficient statistical power to detect program effects.

DISCUSSION

A consistent, accurate method is needed to evaluate the impact of the remarkable scale-up of malaria control interventions in SSA. Clear methods have yet to be identified to directly measure the impact of this scale-up, but this paper proposes an enhanced plausibility design.

The ability to attribute impact to interventions depends on the type of design. The gold standard to attribute causality is

Three analytical plans for validating results from primary analysis

(1) Use of a national platform framework to establish a dose-response relationship between malaria program intensity and malaria health outcomes

Countries may choose to perform a district-level analysis to assess the association between malaria program intensity and outpatient malaria case incidence and inpatient malaria deaths, as reported by routine HMIS. To assess malaria program intensity in the case of vector control, houses sprayed or ITNs distributed per capita per month can be tabulated.⁴⁵ Alternatively, a Bayesian approach can be used to estimate district-level ITN and IRS coverage from nationally representative household surveys.⁴⁷ When conducted appropriately, such an analysis should allow the assessment of the relative contribution of different malaria interventions, above and beyond ecological factors related to transmission risk. During such an analysis, it is imperative to account for contextual factors such as rainfall, temperature, treatment-seeking behaviors, and RHIS reporting quality.

(2) Analysis of cross-sectional datasets to assess the association of ITN exposure (and other control measures) with malaria parasite prevalence and severe anemia and all-cause under-five mortality

Multiple regression modeling is used to assess the association between malaria program exposure (e.g., ITNs) and improvement in malaria-related health outcomes (e.g., reduced prevalence of parasite infections or severe anemia) using population-based household survey data, while accounting for contextual and confounding factors, such as age, sex, mother's education, household socioeconomic status, and geographic location. Survival analysis can also be used to assess the association between household ITN possession and 1–59-month all-cause child mortality based on population-derived cross-sectional data. Such an assessment uses all-cause child mortality as the outcome estimated from a full birth history. Exposure to ITNs in households is ascertained from a net roster of the three years preceding the survey. This analysis allows accounting for potential confounding factors, including calendar year, child's age, and climate (or malaria transmission season), in addition to underlying cluster-level malaria transmission obtained from the Malaria Atlas Project.⁴⁸ Changes in survival should coincide with the timeline of scale-up of malaria control interventions. In both cases for the analysis of cross-sectional data, a matching technique (e.g., propensity score matching) can also be used to help mitigate the bias associated with non-random exposure to malaria control interventions.

(3) Meta-analysis of country results

After impact evaluations have been completed in several countries and comparable datasets are available, a meta-analysis may provide sufficient statistical power and additional insight into malaria program effectiveness at the global level. Meta-analyses can consist of analysis of aggregated point estimates from national-level or subnational studies, typically presented as forest plots, or they can be pooled individual or household-level data, depending on data availability and study objectives. Eisele and colleagues present an example of a pooled individual-level meta-analysis, in which they assessed the association of IPTp with sulfadoxine pyrimethamine and ITNs, or one or the other, with reductions in low birthweight and neonatal mortality, while simultaneously controlling for contextual factors suspected to influence these health outcomes.⁴⁵

FIGURE 3. Three analytical plans for validating results from primary analysis.

an experimental design because of the random assignment of units to intervention and control groups; however, this design is challenging for malaria control programs because intervention areas are based on nonrandom considerations and often are implemented at a national scale. Malaria control programs fall under the category of “complex intervention

research,”⁴⁴ with multiple interventions occurring simultaneously and interacting. In addition, interventions target various groups and sometimes require beneficiaries to adapt a number of behaviors to achieve the intended impact.

The impact evaluation design of malaria control intervention scale-up falls in the area of developing literature on complex or

system-level interventions and plausibility evaluations.^{4,44} This literature suggests that, in addition to measuring changes in key indicators before and after an intervention, an evaluation should include four considerations: 1) a theoretical understanding of how the intervention causes change through an impact model; 2) allowance for variability in outcomes, depending on population characteristics, through the estimation of separate outcomes for different strata; 3) an assessment of the implementation process to determine if any lack of impact reflects implementation failure or genuine ineffectiveness, which can guide and improve program management; and 4) a thorough documentation of context to inform assessments about causality and help generalize results to other contexts.

The improved plausibility design proposed here builds on these principles. The design is advantageous in certain settings because it obviates the need for a control group, and it can assess the impact of national coverage programs. The success of this design, however, requires comprehensive and consistently high-quality data, as well as rigorous analytical techniques. Data from multiple sources in each country should be used to compensate for the shortcoming of any one method of data collection. For example, HMIS often underestimate the total deaths attributed to malaria because most malaria-related deaths in SSA occur outside of public health facilities; however, these data can be combined with confirmed-case reporting to more rigorously assess malaria control intervention impact.⁴⁰ Potential confounders should always be thoroughly investigated in trend analyses and other statistical models to evaluate the role of these factors on the health outcomes of interest.

Although the expanded evaluation framework provides a good basis for evaluating malaria control intervention impact, it does have some limitations. It is still unclear when or at what threshold of intervention coverage an impact evaluation becomes appropriate. The plausibility study design assumes that scale-up and its impact, along with contextual factors, can all be measured reliably at national and subnational levels; however, significant gaps still exist in data acquisition systems in most malaria-endemic countries. The proposed framework relies exclusively on secondary data and its application depends on the quality of the data. Therefore, additional efforts need to be put into strict data collection at the national and subnational levels.

Moreover, baseline data preceding the scale-up of malaria control interventions do not exist for many countries, making it difficult to measure change over time. Countries should continue to invest in strengthening routine health information system to ensure that periodic household surveys are complemented by robust longitudinal data from health facilities where malaria is diagnosed and treated. This is particularly important in changing malaria epidemiological context with countries moving from control to pre-elimination phase. As a starting point, a comprehensive mapping and cataloguing of routine health information systems data elements, data quality and data gaps will be helpful and serve as a baseline for improvement of the systems. Although similar mapping and catalogue exist for household surveys it is lacking for routine health information systems. Furthermore, baseline coverage levels will continue to shift as ongoing scale-up and maintenance of interventions affect malaria transmission and

environmental changes alter malaria intensity. This new context will require changes in evaluation approaches and more reliable, flexible, timely, and representative systems for measuring changes in malaria control interventions and health outcomes.

It is also difficult to develop an accurate counterfactual scenario that captures a prescale-up level of intervention coverage and program intensity. If malaria was already in decline before scale-up, then continuation of prescale-up trends is a conservative or pessimistic counterfactual because this scenario must also assume that baseline coverage increases are also continuing. The program's contribution is then only the recent acceleration of coverage increases, without accounting for the past and current scale-up in coverage. On the other hand, if malaria had been on the rise before scale-up, assuming a continuation of that rise as counterfactual results in an optimistic estimation of impact. Malaria is an epidemic disease, and often a rise is followed by a decline, even without program effort. In addition, urbanization and socioeconomic development may be leading to declining secular trends in most countries.

The relationship between malaria coverage and impact has a time lag, and the dynamics of how burden trends respond to coverage trends is highly nonlinear. Malaria interventions may reach their full impact several years after coverage scale-up; however, that impact may include a partial rebound: as interventions reduce cases, people build up less acquired immunity, which can result in older age groups having a final new stabilized burden level that may not be as low as previous burden levels (e.g., the level 2 years after reaching maximum intervention coverage).

Despite all these challenges, the plausibility approach proposed in this paper provides a good avenue to document evidence of malaria control achievements in endemic countries. However, as data become more available and information systems are strengthened, further improvement to this approach will be needed.

Received May 15, 2015. Accepted for publication November 23, 2015.

Acknowledgments: This work was supported by the President's Malaria Initiative (PMI) through the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with ICF; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of PMI, USAID, CDC or the United States Government.

Disclaimer: Any opinions expressed in this supplement do not necessarily reflect the views of PMI, the United States Agency for International Development, or the Centers for Disease Control and Prevention.

Authors' addresses: Yazoume Ye and Jui A. Shah, MEASURE Evaluation, ICF, Rockville, MD, E-mails: yazoume.ye@icf.com and jui.shah@icf.com. Thomas P. Eisele, Center for Applied Malaria Research and Evaluation, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, E-mail: teisele@tulane.edu. Erin Eckert and Christine L. Hershey, President's Malaria Initiative, United States Agency for International Development, Bureau for Global Health, Arlington, VA, E-mails: eeckert@usaid.gov and chershey@usaid.gov. Eline Korenromp, *Avenir Health* Geneva 2, Switzerland, E-mail: ekorenromp@AvenirHealth.org. Elizabeth Ivanovich, United Nations Foundation, Washington, DC, E-mail: eivanovich@unfoundation.org. Holly Newby, Independent consultant based on Stockholm, Sweden, E-mail: holly.newby@gmail.com. Liliana Carvajal-Velez, Division of

Data, Research, and Policy, Data and Analytics Section, United Nations Children's Fund, New York, NY, E-mail: Icarvajal@unicef.org. Michael Lynch and Achuyt Bhattarai, President's Malaria Initiative, Malaria Branch, Division of Parasitic Diseases and Malaria, Centers for Disease Control and Prevention, Atlanta GA, E-mails: wzl4@cdc.gov and hij6@cdc.gov. Richard E. Cibulskis, Global Malaria Programme, World Health Organization, Geneva, Switzerland, E-mail: cibulskis@who.int. Ryuichi Komatsu, The Global Fund to Fight AIDS, Tuberculosis and Malaria, Zhuzhi Moore, IBTCI, Vienna, VA, E-mail: zhuzhimooore@yahoo.com.

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

REFERENCES

1. WHO, 2014. *World Malaria Report*. Geneva, Switzerland: World Health Organization.
2. Division of Malaria Control, Ministry of Health, Republic of Kenya, *National Malaria Strategy 2001–2010*, April 2001, Nairobi, Kenya.
3. Programme National de Lutte contre le Paludisme, Ministère de la Santé Publique, 2012. *Le plan stratégique de lutte contre le paludisme Madagascar 2013–2017*. Antananarivo, Madagascar.
4. Craig P, Dieppe P, Macintyre S, Michi S, Nazareth I, Petticrew M, 2008. *Developing and evaluating complex interventions: new guidance*. Medical Research Council. Available at: <http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance>. Accessed April 23, 2015.
5. Khandker RS, Koolwal GB, Samad HA, 2010. *Handbook on Impact Evaluation: Quantitative Methods and Practices (World Bank Training Series)*. Washington, DC: The International Bank for Reconstruction and Development, The World Bank. eISBN: 978-0-8213-8029-1.
6. Rowe AK, Steketee RW, Arnold F, Wardlaw T, Basu S, Bakyaite N, Lama M, Winston CA, Lynch M, Cibulskis RE, Shibuya K, Ratcliffe AA, Nahlen BL, Roll Back Malaria Monitoring and Evaluation Reference Group, 2007. Viewpoint: evaluating the impact of malaria control efforts on mortality in sub-Saharan Africa. *Trop Med Int Health* 12: 1524–1539.
7. Habicht JP, Victoria CG, Vaughan JP, 1999. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol* 28: 10–18.
8. Tanzania Malaria Impact Evaluation Research Group, 2012. *Evaluation of the Impact of Malaria Interventions on Mortality in Children in Mainland Tanzania*. Available at http://www.pmi.gov/docs/default-source/default-document-library/tools-curricula/tanzania_ie_report.pdf?sfvrsn=4. Accessed April 23, 2015.
9. Kleinschmidt I, Schwabe C, Benavente L, Torrez M, Ridl FC, Segura JL, Ehmer P, Nchama GN, 2009. Marked increase in child survival after four years of intensive malaria control. *Am J Trop Med Hyg* 2009 80: 882–888.
10. Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ, 2011. *Impact Evaluation in Practice*. Washington, DC: The International Bank for Reconstruction and Development, The World Bank. eISBN: 978-0-8213-8593-7.
11. Victora CG, Schellenberg JA, Huicho L, Amaral J, El Arifeen S, Pariyo G, Manzi F, Scherpbier RW, Bryce J, Habicht JP, 2005. Context matters: interpreting impact findings in child survival evaluations. *Health Policy Plan* 20 (Suppl 1): i18–i31.
12. Lengeler C, 2004. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev* CD000363.
13. Florey LS, Bennett A, Hershey CL, Bhattarai A, Nielsen C, Ali D, Luhanga M, Cameron Taylor C, Eisele T, Ye Y, 2017. Impact of scale-up of district-level insecticide-treated net (ITN) distribution on all-cause child mortality in Malawi, 2004–2010: a summary of regression analysis results. *Am J Trop Med Hyg* 97 (Suppl 3): 65–75.
14. Lindblade KA, Eisele TP, Gimnig JE, Alaii JA, Odhiambo F, ter Kuile FO, Hawley WA, Wannemuehler KA, Phillips-Howard PA, Rosen DH, Nahlen BL, Terlouw DJ, Adazu K, Vulule JM, Slutsker L, 2004. Sustainability of reductions in malaria transmission and infant mortality in western Kenya with use of insecticide-treated bednets: 4 to 6 years of follow-up. *JAMA* 291: 2571–2580.
15. Hershey CL, Bhattarai A, Florey LS, Nielsen CF, Ye Y, Eckert E, Franca-Koh AC, Shargie E5, Komatsu R, Smithson P, Taylor C, Herrera S, Shah J, Mouzin E, Yoon S, Salgado, 2017. Implementing impact evaluations of malaria control interventions: process, lessons learned and recommendations from RBM impact evaluations. *Am J Trop Med Hyg* 97 (Suppl 3): 20–31.
16. World Health Organization, 2010. *A manual for Reviewing the Performance of Malaria Control and Elimination Programmes*, Geneva. Available at: <http://www.who.int/malaria/publications/atoz/whomprmalariaprogramperformancemanual.pdf?ua=1>. Accessed July 21, 2015.
17. Roll Back Malaria, 2013. *Household Survey Indicators for Malaria Control*. Rockville, MD.
18. Ye Y, Wamukoya M, Ezeh A, Emina J, Osman S, 2012. Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Saharan Africa? *BMC Public Health* 12: 741.
19. Korenromp EL, Armstrong-Schellenberg JR, Williams BG, Nahlen BL, Snow RW, 2004. Impact of malaria control on childhood anaemia in Africa—a quantitative review. *Trop Med Int Health* 9: 1050–1065.
20. Onori E, Grab B, 1980. Indicators for the forecasting of malaria epidemics. *Bull World Health Organ* 58: 91–98.
21. Shililu J, Ghebremeskel T, Mengistu S, Fekadu H, Zerom M, Mbogo C, Githure J, Novak R, Brantly E, Beier JC, 2003. High seasonal variation in entomologic inoculation rates in Eritrea, a semi-arid region of unstable malaria in Africa. *Am J Trop Med Hyg* 69: 607–613.
22. Garenne M, Fontaine O, 2006. Potential and limits of verbal autopsies. *Bull World Health Organ* 84: 164.
23. Byass P, 2014. Usefulness of the Population Health Metrics Research Consortium gold standard verbal autopsy data for general verbal autopsy methods. *BMC Med* 12: 23.
24. Soleman N, Chandramohan D, Shibuya K, 2006. Verbal autopsy: current practices and challenges. *Bull World Health Organ* 84: 239–245.
25. Baiden F, Bawah A, Biai S, Binka F, Boerma T, Byass P, Chandramohan D, Chatterji S, Engmann C, Greet D, Jakob R, Kahn K, Kunii O, Lopez AD, Murray CJ, Nahlen B, Rao C, Sankoh O, Setel PW, Shibuya K, Soleman N, Wright L, Yang G, 2007. Setting international standards for verbal autopsy. *Bull World Health Organ* 85: 570–571.
26. Joshi R, Kengne AP, Neal B, 2005. Methodological trends in studies based on verbal autopsies in a rural area of South Africa. *Trop Med Int Health* 5: 824–831.
27. Murray CJL, James SL, Birbaum JK, Freeman MK, Lozano R, Lopez AD, and the Population Health Metrics Research Consortium (PHMRC), 2011. Simplified symptom pattern method for verbal autopsy analysis: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr* 9.
28. Streatfield PK, Khan WA, Bhuiya A, Hanifi SM, Alam N, Diboulo E, Sié A, Yé M, Compaoré Y, Soura AB, Bonfoh B, Jaeger F, Ngoran EK, Utzinger J, Melaku YA, Mulugeta A, Weldearegawi B, Gomez P, Jasseh M, Hodgson A, Oduro A, Welaga P, Williams J, Awini E, Binka FN, Gyapong M, Kant S, Misra P, Srivastava R, Chaudhary B, Juvekar S, Wahab A, Wilopo S, Bauni E, Mochamah G, Ndila C, Williams TN, Desai M, Hamel MJ, Lindblade KA, Odhiambo FO, Slutsker L, Ezeh A, Kyobutungi C, Wamukoya M, Delaunay V, Diallo A, Douillot L, Sokhna C, Gómez-Olivé FX, Kabudula CW, Mee P, Herbst K, Mossong J, Chuc NT, Arthur SS, Sankoh OA, Tanner M, Byass P, 2014. Malaria mortality in Africa and Asia: evidence from INDEPTH health and demographic surveillance system sites. *Glob Health Action* 7: 25369.
29. Phillips DE, AbouZahr C, Lopez AD, Mikkelsen L, de Savigny D, Lozano R, Wilmoth J, Setel PW, 2015. Are well functioning civil registration and vital statistics systems associated with better health outcomes *The Lancet* 386: 1386–1394.
30. Mikkelsen L, Phillips DE, AbouZahr C, Setel PW, de Savigny D, Lozano R, Lopez AD, 2015. A global assessment of civil

- registration and vital statistics systems: monitoring data quality and progress. *The Lancet* 386: 1395–1406.
31. Snow RW, Korenromp EL, Gouws E, 2004. Paediatric mortality in Africa: *Plasmodium falciparum* malaria as a cause or risk? *Am J Trop Med Hyg* 71(Suppl 2): 16–24.
 32. World Health Organization, 2012. *Disease Surveillance for Malaria Elimination: Operational Manual*. Geneva, Switzerland: WHO. ISBN 978 92 4 150333 4.
 33. The Global Fund to Fight AIDS Tuberculosis and Malaria, 2011. *The Global Fund Results Report 2011: Making a Difference*. Geneva, Switzerland: WHO.
 34. Bhattarai A, Ali AS, Kachur SP, Mårtensson A, Abbas AK, Khatib R, Al-Mafazy AW, Ramsan M, Rotllant G, Gerstenmaier JF, Molteni F, Abdulla S, Montgomery SM, Kaneko A, Björkman A, 2007. Impact of artemisinin-based combination therapy and insecticide-treated nets on malaria burden in Zanzibar. *PLoS Med* 4: e309.
 35. Aregawi MW, Ali AS, Al-mafazy AW, Molteni F, Katikiti S, Warsame M, Njau RJ, Komatsu R, Korenromp E, Hosseini M, Low-Beer D, Bjorkman A, D'Alessandro U, Coosemans M, Otten M, 2011. Reductions in malaria and anemia case and death burden to hospitals following scale-up of malaria control in Zanzibar, 1999–2008. *Malaria J* 10: 46.
 36. Kleinschmidt I, Schwabe C, Benavente L, Torrez M, Ridl FC, Segura JL, Ehmer P, Nchama GN, 2009. Marked increase in child survival after four years of intensive malaria control. *Am J Trop Med Hyg* 80: 882–888.
 37. Greenwell F, Neuman M, 2006. *Children's Anemia Levels in West Africa: A Good Proxy for Malaria Morbidity? Population Association of America Annual Meeting*. Los Angeles, California.
 38. Snow RW, Marsh K, 2002. The consequences of reducing transmission of *Plasmodium falciparum* in Africa. *Adv Parasitol* 52: 235–264.
 39. Carneiro I, Roca-Feltrre A, Griffin JT, Smith L, Tanner M, Armstrong Schellenberg J, Greenwood BM, Schellenberg D, 2010. Age-patterns of malaria vary with severity, transmission intensity and seasonality in sub-Saharan Africa: a systematic review and pooled analysis. *PLoS One* 5: e8988.
 40. Roca-Feltrre A, Carneiro I, Smith L, Schellenberg J, Greenwood B, Schellenberg D, 2010. The age patterns of severe malaria syndromes in sub-Saharan Africa across a range of transmission intensities and seasonality settings. *Malaria J* 9: 282.
 41. Bennett A, Yukich J, Miller JM, Vounatsou P, Hamainza B, Ingwe MM, Moonga HB, Kamuliwo M, Keating J, Smith TA, Steketee RW, Eisele TP, 2014. A methodological framework for the improved use of routine health system data to evaluate national malaria control programs: evidence from Zambia. *Popul Health Metr* 12: 30.
 42. Bryce J, Gilroy K, Jones G, Hazel E, Black RE, Victora CG, 2010. The accelerated child survival and development programme in West Africa: a retrospective evaluation. *Lancet* 375: 572–82.
 43. Graves PM, Osgood DE, Thomson MC, Sereke K, Araia A, Zerom M, Ceccato P, Bell M, Del Corral J, Ghebreselassie S, Brantly EP, Ghebremeskel T, 2008. Effectiveness of malaria control during changing climate conditions in Eritrea, 1998–2003. *Trop Med Int Health* 13: 218–228.
 44. de Savigny D and Adam T, eds., 2009. *Systems Thinking for Health Systems Strengthening: Alliance for Health Policy and Systems Research and WHO*. Geneva, Switzerland.