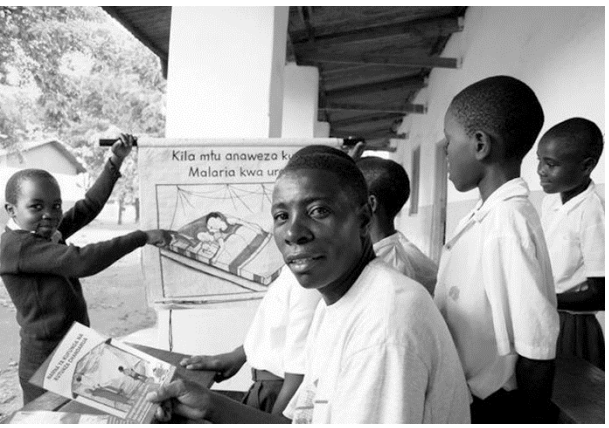
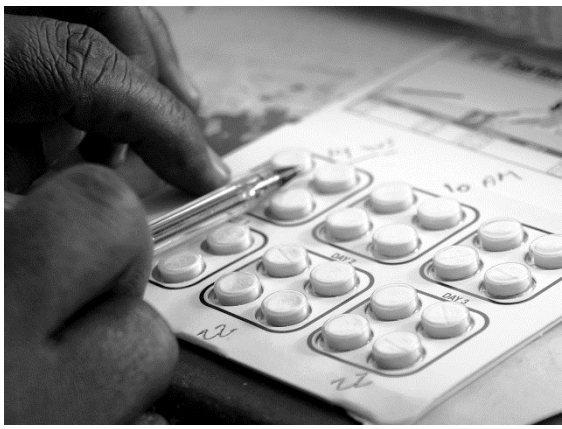
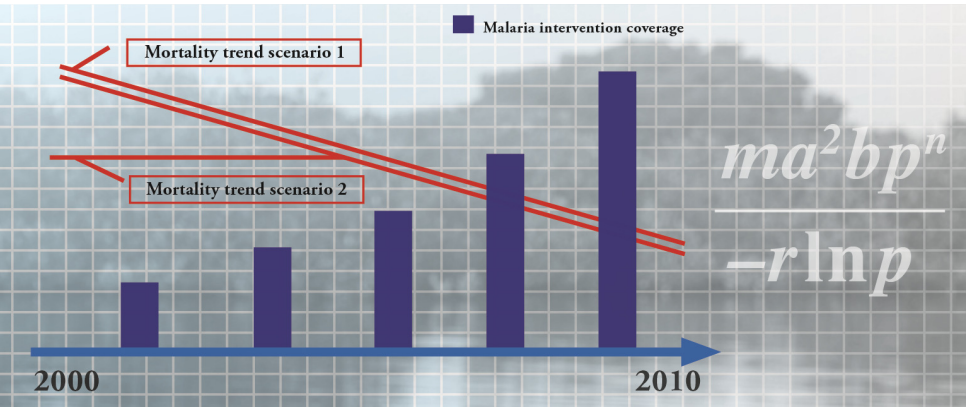






Guidance for Evaluating the Impact of National Malaria Control Programs in Highly Endemic Countries

December 2014

CDC
USAID
FROM THE AMERICAN PEOPLE
President's Malaria Initiative
MEASURE Evaluation
ROLL BACK MALARIA

GUIDANCE FOR EVALUATING THE IMPACT OF NATIONAL MALARIA CONTROL PROGRAMS IN HIGHLY ENDEMIC COUNTRIES

Contents

ACKNOWLEDGMENTS **1**

GLOSSARY OF KEY TERMS **2**

ABBREVIATIONS **4**

I. EXECUTIVE SUMMARY **5**

II. INTRODUCTION AND OBJECTIVES **8**

II.1 The Problem 8
II.2 Immediate Needs for Consistent Measurement 9
II.3 Objectives of the Framework Document 10
II.4 Target Audiences 10
II.5 Limitations 11

III. IMPLEMENTING IMPACT EVALUATIONS **12**

III.1 Process 12
III.1.1 Engage Stakeholders 13
III.1.2 Describe Malaria Control Program 13
III.1.3 Design an Evaluation 14
III.1.4 Generate Credible Evidence 15
III.1.5 Promote Use and Disseminate Evaluation Findings 15
III.2 Challenges 16
III.2.1 Different Needs for the Evaluation 16
III.2.2 Data Access 16
III.3 Description of Resources 17
III.3.1 Timeline 17
III.3.2 Human Resources and Skills 18
III.3.3 Financial Resources 18

IV. EVALUATION DESIGN **20**

IV.1 Aim of the Impact Evaluation in Each Country 20
IV.2 Constraints in Evaluating Malaria Programs 20
IV.3 Study Design Options 21
IV.4 Plausibility Design 23
IV.5 What Results Might Look Like 25

V. PROGRAM DESCRIPTION **28**

V.1 Country Background 28
V.2 Health Care System 28
V.3 Malaria Epidemiologic Profile 29
V.4 Malaria Control Strategy 29
V.5 Financial Resources 30
V.6 Commodity Inputs 31
V.7 Key Events 32

VI. MEASURING THE COVERAGE OF MALARIA CONTROL INTERVENTIONS **33**

VI.1 Outcome Indicators Derived from Population-based Household Surveys 35
VI.1.1 Considerations 35
VI.2 Outcome Indicators from Routine Health Information Systems 36
VI.2.1 Considerations 37

VII. MEASURING TRANSMISSION INTENSITY 39
VII.1 Transmission Intensity 39
VIII. MEASURING MALARIA MORBIDITY 41
VIII.1 Population-based Indicators 41
VIII.2 Health Facility-based Indicators 43
IX. MEASURING MORTALITY 46
IX.1 Malaria-specific Mortality 46
IX.1.1 Routine Health Management and Information Systems and Sentinel Surveillance 46
IX.1.2 Health and Demographic Surveillance Systems 50
IX.1.3 Demographic and Health Surveys and National Censuses 52
IX.1.4 Ways Forward—Examine, Improve Data Quality and Completeness 52
IX.2 All-Cause Child Mortality from Household Surveys 53
X. MEASURING AND ACCOUNTING FOR NONMALARIA PROGRAMS, CONTEXTUAL FACTORS 59
X.1 Data Sources 60
X.2 Accounting for Environmental Factors 61
X.2.1 Temperature 61
X.2.2 Rainfall and Humidity 63
X.2.3 Land Cover 65
X.3 Data Analysis 65
XI. DATA SYNTHESIS, TRIANGULATION, AND INTERPRETATION 68
XI.1 Trends Over Time 68
XI.2 Further Analysis 69
XI.3 Multicountry Meta-analysis 70
XI.4 Synthesis of Results 70
APPENDICES 73
Appendix A. Limitations and Assumptions of the Plausibility Study Design 74
Appendix B. What Results Might Look Like: Example Scenarios 75
Appendix C. Malaria Risk Parameters 77
Appendix D. Details of Analyses to Help Validate and Quantify Results Obtained from the Plausibility Study Design 78
Appendix E. Country Case Studies on the Analytic Approaches 81
Appendix F. Case Study 1: Bioko Island 83
Appendix G. Case Study 2: Zambia 85
Appendix H. Case Study 3: Zanzibar, North A District 87
Appendix I. Case Study 4: Mainland Tanzania 89
Appendix J. Case Study 5: Malawi 92

Acknowledgments

This document was initiated by the Mortality Task Force of Roll Back Malaria's Monitoring and Evaluation Reference Group. The following task force members contributed as authors of this document:

Yazoume Ye (MEASURE Evaluation/ICF International), Thom Eisele (Tulane University School of Public Health and Tropical Medicine), Jui A. Shah (MEASURE Evaluation/ICF International), Eline Korenromp (Futures Institute), Christie Hershey (President's Malaria Initiative/United States Agency for International Development), Elizabeth Ivanovich (United Nations Foundation), Holly Newby (United Nations Children's Fund), Liliana Carvajal-Velez (United Nations Children's Fund), Mike Lynch (World Health Organization), Ryuichi Komatsu (The Global Fund to Fight AIDS, Tuberculosis and Malaria), Richard Cibulskis (World Health Organization), Zhuzhi Moore (The DHS Program/ICF International), Achuyt Bhattarai (President's Malaria Initiative/Centers for Disease Control and Prevention), and Erin Eckert (President's Malaria Initiative/United States Agency for International Development).

The authors have tremendous gratitude for those who reviewed this document and provided constructive feedback to ensure that the document is relevant, accurate, and concise: Emmanuela Gakidou, Abdisalan Noor, Alexander Rowe, and Steve Yoon.

Special thanks also go to Lynne Jennrich and Synergy Enterprises, Inc. for editing and proofreading multiple versions of this document and to Elizabeth Notter, Bertha McMaster, and Charlene McNamara for graphics and layout. Cover photos and graphic credits (clockwise from top left): Ayesha Mawji and JHUCCP/USAID COMMIT Project, courtesy of Photoshare; Bonnie Gillespie, courtesy of Photoshare; Arturo Sanabria, courtesy of Photoshare; Arturo Sanabria, courtesy of Photoshare; and Yazoume Ye.

MEASURE Evaluation is funded by the U.S. Agency for International Development (USAID) through cooperative agreement GHA-A-00-08-00003-00. The views expressed in this publication do not necessarily reflect the views of USAID or the United States Government.

Suggested citation: Mortality Task Force of Roll Back Malaria's Monitoring & Evaluation Reference Group. 2014. *Guidance for Evaluating the Impact of National Malaria Control Programs in Highly Endemic Countries*. Rockville, MD, USA: MEASURE Evaluation.

Glossary of Key Terms

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.
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 (World Bank, 2006).
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 (World Health Organization, 2012).
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.
Direct malaria 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."
Endemic malaria	Term used to describe ongoing malaria with a measurable incidence of cases and mosquito-borne transmission in an area over a succession of years (World Health Organization, 2012). Also known as "stable malaria."
Epidemic malaria	Term used to describe occasional malaria outbreaks in normally malaria-free regions; a particularly severe malaria season in a normally low-risk area. Also known as "unstable malaria."
Evaluation	Periodic assessment of whether objectives are being achieved, often requiring special surveys or studies (Gertler, 2011).
Health facility-based malaria morbidity indicators	Indicators of morbidity based on data from surveillance and routine information systems, such as health facility registries or health management information systems. Examples are malaria outpatient visits or cases, hospital inpatient admissions, and outpatient visits and hospitalizations for severe anemia in young children in high-endemic settings.
Indirect malaria mortality	Deaths in which malaria was a contributing cause, and the death was categorized as a non-malaria death. Examples are deaths from the combined effects of malaria-associated anemia and pneumonia, in which the cause was categorized as pneumonia; deaths linked to low birth weight caused by malaria in the mother during pregnancy; or deaths resulting from consequences of clinical management, such as HIV exposure from a blood transfusion for malaria-related anemia or sequelae of a malaria infection, such as epilepsy caused by cerebral malaria (Snow, et al., 2003).
Malaria parasite prevalence	Proportion of children ages 6–59 months with malaria parasite infection (Roll Back Malaria guidelines, 2009).
Malaria transmission	Spread of malaria by completion of a full transmission cycle (man→mosquito→man).
Malaria transmission intensity	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.

Terms	Definition
Malaria-related mortality	Deaths in which malaria was the underlying cause or a contributing cause; the sum of direct and indirect malaria mortality (Rowe, et al., 2007).
Monitoring	Ongoing tracking of progress toward an objective, often using routinely collected data (MEASURE Evaluation).
Parasitemia	Presence of parasites in the blood; number of parasites per volume of blood.
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 (Habicht, et al., 1999 and Morgenstern 1982).
Population-level malaria morbidity indicators	Indicators on malaria morbidity collected through population-based surveys; examples are malaria parasite prevalence and anemia.
Under-5 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 (Garenne & Fontaine 1990; Anker, et al., 1999; Soleman, et al., 2006). To attribute causes of deaths, interviews are analyzed by an algorithm or clinicians who decide on causes by majority vote (Rowe, et al., 2007).

Abbreviations

ACCM	All-cause child mortality
ACT	Artemisinin combination therapies
BIMCP	Bioko Island Malaria Control Project
CRVS	Civil registration and vital statistics
DHS	Demographic and Health Surveys
EIR	Entomological inoculation rate
GDP	Gross domestic product
Global Fund	The Global Fund to Fight AIDS, Tuberculosis and Malaria
HBR	Human biting rate
HDSS	Health and Demographic Surveillance System
HMIS	Health management information system
IGME	Inter-agency Group for Child Mortality Estimation
IPTp	Intermittent preventive treatment for pregnant women
IRS	Indoor residual spraying
ITNs	Insecticide-treated nets
LiST	Lives Saved Tool
LLINs	Long-lasting insecticidal nets
M&E	Monitoring and evaluation
MDGs	Millennium Development Goals
MERG	Monitoring and Evaluation Reference Group
MICS	Multiple Indicator Cluster Surveys
MIS	Malaria Indicator Surveys
MoH	Ministry of Health
NMCC	National Malaria Control Center
NMCP	National Malaria Control Program
***Pf*PR**	*Plasmodium falciparum* parasite rate
PMI	President's Malaria Initiative
RBM	Roll Back Malaria
RDT	Rapid diagnostic test
RHMIS	Routine health management information system
UN	United Nations
UNDP	United Nations Development Programme
UNICEF	United Nations' Children Fund
VA	Verbal autopsy
VC	Vectorial capacity
WHO	World Health Organization

I. Executive Summary

Over the past decade, funding for malaria programs has increased significantly, especially in Sub Saharan Africa. This has led to the scale-up of key interventions such as insecticide-treated nets, indoor residual spraying, intermittent preventive treatment for pregnant women, and treatment. In 2007, the Roll Back Malaria Partnership's Monitoring and Evaluation Reference Group (MERG) proposed the use of a plausibility design to measure impact of malaria control programs. Since then, new measurement needs and evidence have emerged, requiring an updated approach.

This document reviews and updates the 2007 evaluation framework and provides recommendations for evaluating the scale-up of malaria control interventions in endemic countries. The document, which is not intended to be an exhaustive resource on statistical modeling techniques, targets staff of National Malaria Control Programs, Ministries of Health, development partners, and people conversant with monitoring and evaluation. The methods described are most relevant for countries with a high malaria burden, which tend to struggle with capturing malaria deaths and where malaria constitutes a large portion of overall mortality.

The suggested evaluation framework consists of five interdependent steps: (1) engage stakeholders, (2) describe the malaria control program, (3) design the evaluation, (4) generate credible evidence, and (5) promote use and dissemination of findings. To ensure that the evaluation results are relevant and useful in adjusting programs and implementing policies, an evaluation should not be protracted. Evaluators should note that the planning, analysis, writing, and review phases of an impact evaluation require substantial time, in addition to money and staff.

Due to challenges with measuring malaria-specific mortality and defining a control or comparison group required for a traditional impact evaluation, MERG suggests that full-coverage malaria control programs rely on an ecological study design, often referred to as a plausibility study design. This design assesses the simultaneous changes in intervention coverage and the malaria burden at the population level. An important aspect of this design is that the assessed interventions have empirical proof of impact on the outcomes measured. In that regard, if the evaluation showed that intervention coverage increased sufficiently with simultaneous improvements in population-level outcomes (malaria morbidity and mortality), it is plausible that the program contributed to the improvements. Indicators at the end of the evaluation period are compared with the counterfactual, which is the assumption that pre-intervention trends would have continued.

A malaria control program description is developed as part of the early steps of conducting an evaluation. The program description illustrates both how the program was expected to work and what was actually done, complete with a timeline of major efforts. This helps readers understand relevant milestones, policies, and activities that may have led to impact in malaria outcomes.

Evaluations generally use population-based data to measure coverage indicators; however, routine health data might also contribute to understanding intervention coverage. MERG suggests the use of 11 primary outcome indicators to assess the scale-up of key malaria control interventions. The primary sources of data for these indicators are the Demographic and Health Surveys, Malaria Indicator Surveys, and Multiple Indicator Cluster Surveys. Looking at these indicators can help evaluators assess to what extent malaria control interventions were scaled up and whether they reached target populations.

It is also important to document malaria transmission risk. Malaria risk can be measured by using the entomological inoculation rate, parasite prevalence, or parasite incidence in a number of ways. These different measures can be used to compare different geographical settings within a country, analyze different periods of time, or examine trends over time. To understand transmission risk, it is important to have regular data on temperature, which plays a major role in the malaria vector life cycle, and rainfall, which allows for mosquito breeding. Because these factors can vary greatly within a small area, they are particularly useful for subnational analyses.

Reductions in malaria prevalence should have the desired effect of reducing malaria-specific morbidity, which can be measured both at the population and facility levels. Population-based data are particularly useful in areas with high-intensity malaria transmission; in such situations, anemia and malaria parasite prevalence should be measured in children under age 5 years. Morbidity indicators based on health facility data may be calculated among all ages or solely among children under age 5 years. It is also useful to assess laboratory testing coverage, malaria test positivity rate, and proportion of suspected malaria cases that received a laboratory diagnosis when interpreting facility-based data.

It is difficult to measure malaria-specific mortality, partly because of malaria's large contribution to increasing the risk of death from other common illnesses. In countries with a high malaria burden, however, malaria is a key contributor to all-cause child mortality (ACCM). MERG recommends using ACCM as the standard measure of mortality impact of malaria interventions for highly endemic settings. Evaluations can measure malaria deaths from a number of routine data sources, surveys, or national census, each of which has unique strengths and challenges. Combining data collected from multiple methods may yield the strongest estimates of malaria mortality.

Contextual factors have the potential to (1) confound the association between scale-up of the intervention and ACCM and (2) modify the effect of the intervention, and thus, affect the conclusion. Contextual factors vary from country to country, but generally include other child survival interventions, climatic and environmental factors, health systems factors, and socioeconomic factors of the country's population. Evaluations should list relevant contextual factors and assess whether to change them and to what degree they affect ACCM.

The evaluation will bring all of these data together in an analysis of trends over time. Further analyses can help validate positive effects found by the plausibility design. Modeling, using the Lives Saved Tool for example, can help estimate the number of malaria deaths averted, as well as to support results and plan next steps.

The plausibility study design assumes that scale-up and its impact, along with confounding factors, can be reliably measured, but significant gaps exist in data acquisition systems of most malaria-endemic countries. Moreover, baseline data do not exist for many countries, making it difficult to assess impact of interventions. Thus, countries should continue to develop their data acquisition systems and strive to make them more robust and reliable. Furthermore, baseline measures will also continue to shift as the scale-up influences malaria transmission, and environmental changes alter malaria intensity. Thus, countries likely will need to modify and tailor their intervention strategies to account for these changes. In this context, MERG is calling for a more reliable system to measure interventions and health outcomes.

II. Introduction and Objectives

Key points

- Over the last decade, funding for malaria programs has increased significantly, especially in Sub Saharan Africa, which has led to the scale-up of key interventions (insecticide-treated nets, indoor residual spraying, intermittent preventive treatment for pregnant women, and effective treatment).
- The purposes of this guidance document are to (1) review and update the evaluation framework as proposed by the Monitoring and Evaluation Reference Group in 2007 and (2) provide recommendations for evaluating the scale-up of malaria control interventions in highly malaria-endemic countries.
- The target audiences of this guidance document are staff of National Malaria Control Programs Ministries of Health, and funding agencies, in addition to people who work in monitoring and evaluation.
- This guidance document is not intended to be an exhaustive resource on evaluation methodology.

II.1 The Problem

After a decade of increased efforts in malaria control, the burden of malaria remains unacceptably high despite the existence of effective tools for control and proven intervention strategies. The failure of the malaria eradication campaign in the 1950s reduced public interest in malaria control as a communal goal, and funding and resources languished. To strengthen malaria control programs, the Roll Back Malaria (RBM) Partnership in 2000 set an ambitious agenda to reduce by 50% malaria-related mortality by 2010 and then updated its goals to targeting near-zero preventable deaths by 2015. RBM looks to achieve these goals through the scale-up of effective interventions, such as insecticide-treated nets (ITNs), prompt and effective treatment of malaria cases, intermittent preventive treatment for pregnant women (IPTp), and indoor residual spraying (IRS), where applicable. From 2000 through the present day, funding for malaria programs has increased exponentially, and malaria has become a high priority on the international agenda.

With increased resources comes a responsibility to be accountable, learn lessons to further improve the effectiveness of investment, and justify continued and potentially increased commitments. To that end, the Monitoring and Evaluation Reference Group (MERG) of the RBM Partnership has developed specific guidance on evaluation frameworks, indicators, and data collection systems to track progress and show results. In 2007, the MERG proposed an approach to evaluating the impact of malaria control programs in Sub Saharan Africa that involved tracking the increased coverage of scientifically proven interventions and a simultaneous decrease in cause-specific morbidity and all-cause child mortality (ACCM).[1] It lays out a plausibility argument that demonstrates how these

malaria control measures lead to success. MERG outlined many of these methods as they existed seven to eight years ago, but the intervening years have seen much progress on data collection and interpretation of results. In light of measurement needs for the 2015 Millennium Development Goals (MDGs), now is the time to pull together all evidence on the effectiveness of these measurement methods and the appropriate use of each. This report provides an overview of measurement issues for monitoring and evaluating the scale-up of malaria control programs, and assessing the impact of these programs on morbidity and mortality.

II.2 Immediate Needs for Consistent Measurement

The year 2010 was important for RBM and many of its partners in the fight against malaria. In April 2000, the nations of Sub Saharan Africa committed to the goal of halving the burden of malaria mortality by 2010. The resulting Abuja Declaration, signed by delegations from 44 malaria-endemic African nations, combined with further scientific evidence on the efficacy of malaria control interventions, formed the bedrock for malaria scale-up in the past decade. Also, in 2000, the United Nations (UN) countries adopted the MDGs, two of which, (4) Reduce Child Mortality and (6) Combat HIV/AIDS, Malaria, and Other Diseases, relate directly to malaria. In fall 2010, a high-level meeting of the UN General Assembly reaffirmed the UN commitments to achieve these goals, including reductions in child mortality and the burden of malaria across all endemic countries. At the same time, institutional donors and RBM partners, such as the U.S. Government, World Bank, Bill and Melinda Gates Foundation, and The Global Fund to Fight AIDS, Tuberculosis and Malaria, are all undertaking serious evaluations of their efforts in malaria control. In this context, the need is critical for rigorous analytic methods to evaluate the effects of expanded interventions and consistent reliable measures of impact. This document provides this guidance to all partners, so that each can contribute consistently to the larger effort.

This document describes methods to examine cumulative contributions to achieve national and international malaria control goals. The objective is not to attribute or apportion change to any specific intervention or to any specific donor's efforts. The analyses described are not intended to evaluate the efficacy of a specific intervention; many published studies clearly make cases for specific interventions. Rather, this framework looks at the results of collective efforts over a period extending back to 2000, the year of the Abuja Declaration. During that time frame, many interventions have been scaled up; donors have provided assistance, and staff and funds have increased exponentially. This document describes how to take all those inputs into account; control for contextual factors, such as climate, urbanization, and other health-related activities; and estimate the changes in morbidity and mortality over the past decade.

It bears mentioning that such an analysis without attribution is not, in the strictest sense, an impact evaluation. The standard academic definitions of evaluation attribute change in impact measures directly to program interventions. The most rigorous evaluations designate a specific counterfactual (a randomized control group, for example) to control for external factors that affect the outcomes.

However, malaria programs are scaling up interventions that have already been proven scientifically. Scale-up of interventions has been rapid and nationwide over the past decade, or longer, and ethically it would not be possible to withhold lifesaving interventions from a population to gain a control group. Such a narrow definition of "impact evaluation" is less relevant and practical when measuring effectiveness of programs that use proven interventions in field settings. To some extent, however, it is possible to link scale-up of interventions to changes in morbidity and mortality. The outside contextual factors are controlled through the application of regression techniques in the analysis, rather than through the study design. The methods used are as robust as possible within this context to demonstrate program accomplishments over a period of time.

II.3 Objectives of the Framework Document

The main objective of this framework is to describe how to evaluate the impact of malaria control programs in highly endemic countries. The framework seeks to achieve these specific objectives:

1. Update the evaluation design proposed by the MERG in 2007 by taking into account new data and recent experiences in conducting evaluations.
2. Make recommendations for evaluating the scale-up of malaria control interventions in the context of a national malaria control program.
3. Summarize recent experience and data on morbidity and mortality measurement gathered using various methods and data sources.

This document is based on the RBM Partner's Expert Consultation on Mortality Measurement that occurred in April 2010. At that meeting, experts presented recent developments in measuring ACCM and malaria-specific mortality, including estimation procedures. As a result, the idea of this guidance document emerged, and a draft outline was developed. During several months, staff from stakeholder agencies contributed to the development of sections of the document. These sections were reviewed and discussed in bimonthly teleconferences among the agencies. In October 2010, country program staff met in Dar es Salaam, Tanzania, to discuss how to measure the impact of malaria programs, obtain feedback from program managers, and orient the country programs to planned evaluation efforts. Based on multiple consultations, chapters were added to cover transmission intensity and the evaluation implementation. Feedback and responses from country programs, as well as expert review, informed the final document.

II.4 Target Audiences

The audience of this guidance document is the staffs of National Malaria Control Program (NMCP) offices, Ministries of Health (MoH), and donor agencies interested in evaluating the impact of the scale-up of national malaria control programs. It is for people conversant with monitoring and evaluation (M&E) principles and tools. Although the guide presents and discusses methods for evaluation, including their strengths and weaknesses, it is not intended to be an exhaustive resource

on evaluation methodology. This information appears in other documents, and, where appropriate, we provide relevant references.

II.5 Limitations

This document is based on the RBM Expert's Consultation on Mortality Measurement meeting, subsequent meetings (including the Multiagency Impact Evaluation Workshop), and discussions with RBM partner agencies throughout 2011. It reflects the current consensus of group members on measurement methods and techniques and their appropriate use. The authors recognize that malaria epidemiology and control are changing rapidly, in response to the scale-up of malaria control interventions, and to factors such as urbanization, climate change, and insecticide resistance. This document is intended to be a living document that is current and accurate at the time of publication, but subject to updates and revisions as global needs for measurement change.

Sub Saharan Africa bears the greatest malaria burden. Consequently, this document focuses on measurement relevant to Africa and highlights data collection tools, such as the Malaria Indicator Surveys (MIS), and analytic approaches now used there. Changes might be needed in the coming years, as epidemiology of the disease shifts. In addition, after much discussion, the authors decided Asia and Central and South America merit their own guidance documents, which could focus on tools and methods commonly used in those regions, with greater emphasis on case detection and surveillance, for example.

---

[1] Rowe AK, Steketee RW, Arnold F, Wardlaw T, Basu S, Bakyaita N, et al. Viewpoint: Evaluating the impact of malaria control efforts on mortality in sub-Saharan Africa. *Trop Med Int Health*. 2007. 12(12):1524–1539.

III. Implementing Impact Evaluations

Key Points

- A framework for impact evaluation consists of five interdependent steps: (1) engage stakeholders, (2) describe the malaria control program, (3) design an evaluation, (4) generate credible evidence, and (5) promote use and disseminate findings.
- Stakeholder engagement is essential throughout the impact evaluation to ensure credibility, transparency, and, ultimately, use of the evaluation findings.
- It is advisable to involve a local partner to coordinate the impact evaluation in country and hire additional staff, including an analyst with epidemiology and biostatistics skills, as needed.

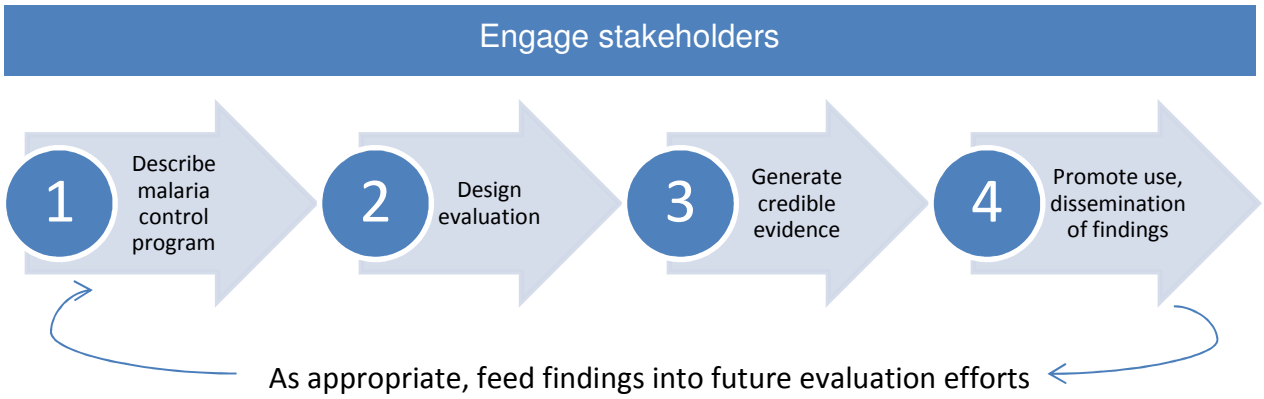
III.1 Process

The following framework for implementing an impact evaluation applies, regardless of what kind of organization leads the evaluation. An external organization—such as funding partners, technical organization, or academic institution—or an in-country organization, such as National Malaria Control Program (NMCP), Ministry of Health (MoH), or academic institution, may lead an evaluation. Even if an organization external to the country leads the evaluation, the NMCP and MoH should lead the process in country and have ownership.

International agencies that seek to evaluate the impact of overall malaria activities in a given country should align their evaluation needs and collaborate on a single evaluation report. This will reduce the reporting burden on country programs. International agencies are encouraged to review recent evaluations before launching a new one.

Figure III.1 shows the five interdependent steps of an evaluation: (1) engage stakeholders, (2) describe the malaria control program, (3) design an evaluation, (4) generate credible evidence, and (5) promote use and dissemination of findings. Evaluators should finish each step before proceeding to the next one, but during the complex evaluation process, the steps may repeat.

FIGURE III.1: IMPLEMENTATION FRAMEWORK FOR EVALUATING THE EFFECT OF MALARIA CONTROL PROGRAMS



III.1.1 Engage Stakeholders

Stakeholder involvement is central to all evaluations. The term "stakeholder" refers to "individuals, groups, or organizations that have an interest in a program. Stakeholders may include funding agencies, policymakers, planners, advocacy groups, communities, or groups that the program is intended to benefit, and other groups that might be affected negatively or positively."[1] Stakeholder engagement is the first step toward establishing the need for an evaluation and promoting ownership of the evaluation findings.[2] Identifying evaluation stakeholders may not be an entirely new process because stakeholders engaged during the malaria strategic planning remain relevant to the evaluation. To ensure a credible, efficient evaluation, it is critical to strike a balance between stakeholder size and engagement level. Use of standard stakeholder analysis can guide the choice of stakeholders in the evaluation.[3,4] It is essential to engage a broad range of stakeholders.

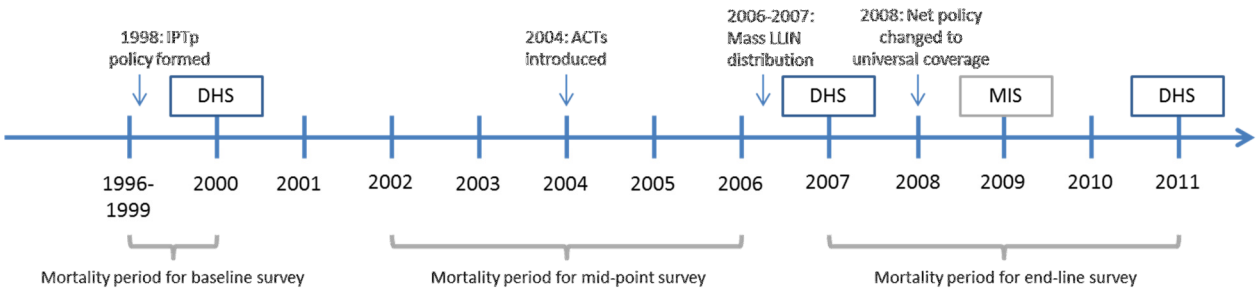
Evaluators should engage stakeholders from the beginning and consult them throughout the process. NMCP and MoH have important leadership roles in convening and engaging stakeholders in the evaluation. Existing mechanisms, such as monitoring and evaluation steering committees or NMCP or MoH workings groups—rather than duplicate or parallel mechanisms—are best suited to engage stakeholders.

III.1.2 Describe Malaria Control Program

The second step in the evaluation is to describe a malaria control program to illustrate the logic of its implementation. This description includes control strategies and their rationale, implementation plans, outputs and expected outcomes, and their interrelationships. The program description provides a clear indication of when and where things happened during the period of evaluation. It is also useful to put current malaria control activities in the context of historical malaria control efforts.

This often is illustrated graphically as a logic or impact model, which is a critical element in the evaluation implementation framework.[5,6,7,8]

FIGURE III.2: EVALUATION TIME FRAME



Notes: The mortality trend is measured with three surveys that each give a five-year mortality estimate: a 2000 baseline survey provides an estimate for 1996–2000, a 2006 mid-point survey provides an estimate for 2002–2006, and a 2011 end-line survey provides an estimate for 2007–2011. Also shown are intervention rollout and policy changes.

Source: Adapted from "Evaluation of the Impact of Malaria Control Interventions on All-Cause Mortality in Children Under Five in Uganda." Draft documents, March 2013.

III.1.3 Design an Evaluation

This step comprises the following key elements: evaluation objectives, questions, methods, and time frame. Input from stakeholders should guide development of the evaluation objectives and questions. Evaluation objectives should be realistic for the period covered. The key evaluation questions,[9] often comprise descriptive, normative, and cause-and-effect questions. The questions should align with evaluation objectives and be answerable within the time, resources, and data available. In developing evaluation questions, it is important to consider what stakeholders will need to know to use the information and results of the evaluation.

Evaluators should clearly describe the methods used. They should choose methods and design analyses to address the evaluation questions based on an assessment of available data sources.

Evaluators should define an evaluation time frame with an appropriate baseline and end line. Countries should define these evaluation periods based on when funding disbursements and malaria control interventions occurred. In addition to defining the evaluation periods, each country timeline should clearly illustrate the period when and where malaria control interventions were deployed, surveys were conducted, and the periods encompassed by mortality estimates, as well as other relevant information (see Figure III.2 for an example). After the design is complete, all stakeholders can help develop and agree to a detailed analysis plan, which should include a tabulation plan. For each hypothesis or question being asked, the design should outline the data sources needed, methods to be used, and analyses to be conducted.

III.1.4 Generate Credible Evidence

Compiling information that stakeholders perceive as trustworthy and relevant for answering the evaluation questions improves evaluation credibility. Aspects of generating evidence include defining evaluation indicators, identifying data sources and assessing their quality, analyzing the data, and writing and reviewing the report.

The recommended outcome (malaria control intervention coverage) and impact (malaria morbidity and mortality) indicators to consider for an evaluation of the impact of malaria control programs appear in Table VI.1. Evaluators should identify appropriate and available data sources to measure these indicators through literature review, discussion with experts, and stakeholder meetings. National evaluations of the type proposed here rely exclusively on secondary analysis of existing data that may have been collected for other purposes. Obtaining these datasets may require a memorandum of understanding (MoU) or data-use agreement. If access to some datasets is not possible for the evaluation team, several options exist. A local partner with access to the datasets could complete an analysis and share results with other evaluation partners. As an alternative, when evaluation partners cannot obtain original datasets, evaluators could abstract values for the indicators of interest and their standard errors from published or unpublished reports of surveys and studies.

During the initial steps of evaluation planning, all stakeholders should agree on who will analyze each of the available data sources and who will write each section of the report. Multiple parties can write sections of the report, but the group should select one person to compile the report.

Stakeholders should agree on a review process and timeline at the outset. Because the report may go through several drafts, partners should identify someone or several people to review each draft and agree on a timeline for reviews. Each partner involved in the evaluation should have an opportunity to review the report. When external partners are conducting an evaluation, reviewers should include people who know the country context to help interpret the results. A follow-up stakeholders' consultation can provide a forum for presenting evaluation results and seeking feedback from people with a broad range of perspectives. At the outset, stakeholders should also agree on the process for final approval of the evaluation report. Typically, the NMCP and primary partners conducting the evaluation should approve the final product. It is important to factor in the time required for approval by NMCP and different agencies.

III.1.5 Promote Use and Disseminate Evaluation Findings

The following five steps can help promote the use of evaluation findings: (1) design, (2) preparation, (3) feedback, (4) follow-up, and (5) dissemination.[10] Involving stakeholders in the previous four steps of the evaluation will also help promote use of the findings. Stakeholders should play a role in the planning, evaluation design, analysis, and report writing. Gaining buy-in and support from senior officials of the host-country government will help increase the reach of evaluation findings.

To ensure the best use of data collected, stakeholder feedback and information gathered during follow-up meetings should form the basis of the final product. During initial stakeholder meetings, partners should decide on the format of the final product (electronic format only or hard-copy reports). The results can also be disseminated to relevant audiences through other means, including peer-reviewed publications, policy briefs, factsheets, and workshops.

Results are intended to help multiple stakeholders. NMCPs can use evaluation findings to inform the development of national malaria control strategic plans. Financial partners can use the findings to demonstrate their contributions to malaria control in the given country and guide further funding decisions. In addition, limitations and gaps in data availability identified during the evaluation should be used to guide national monitoring and evaluation (M&E) strategies and plans, and inform donor funding decisions for future M&E.

III.2 Challenges

Several challenges can arise during multiagency impact evaluations, some of which are discussed here with suggestions to address them.

III.2.1 Different Needs for the Evaluation

NMCP needs and desired timing for an evaluation may not match those of financial partners and international agencies. Understanding the needs of each stakeholder will help evaluators agree on a way forward. Reasons for requesting an evaluation could range from updating the malaria control strategic plans to accounting for donor funds and renewing grants. International agencies should make every attempt to align their evaluation needs and collaborate on a single evaluation report. Initial stakeholder meetings also should promote buy-in, address partners' concerns, and result in an evaluation design that meets the needs of all partners involved.

Where a joint evaluation is not possible, for example because of differences in evaluation approaches or strict timing requirements, the separate evaluation teams should still attempt to coordinate. This coordination can help teams spot different results and explore potential explanations for these differences to avoid confusion at the country level.

III.2.2 Data Access

Data access can present a challenge. It is important to identify the primary owner of data sources because not all datasets will be available publicly. It may take time to gain access to the datasets, and therefore, it is important to request access as early as possible in the planning stages of the evaluation. As mentioned earlier, an MoU or data-use agreement may be necessary to outline exactly how the data will be used and who will have access to it. This also may require partners to agree if findings obtained from the evaluation will be used only in the evaluation report or if they also may be used in journal articles derived from the evaluation.

III.3 Description of Resources

III.3.1 Timeline

To ensure that the evaluation results are relevant and useful in program adjustments and policy implementation, an evaluation should not have a protracted timeline. However, it will require substantial time for the planning, analysis, writing, and review phases. Approximate times needed for the evaluation phases appear in Figure III.3.

Several actions taken early in the evaluation planning can help ensure smooth completion of the report. First, identify relevant stakeholders, especially those who will be involved in the analysis and writing of the report; seek more staff if necessary. Second, gain access to the datasets. Third, allocate sufficient time for data analysis, report writing, review, and dissemination. Deadlines should be set and adhered to for each draft of the report.

FIGURE III.3: ILLUSTRATIVE TIMELINE FOR CONDUCTING AN IMPACT EVALUATION

Activity	Estimated time	Parties involved
Start discussion with in-country and international stakeholders	4 weeks	Funding partners, national authorities
Develop work plan, analysis plan, and task matrix	2 weeks	Evaluation team
Identify and contract implementation partner and any remaining members of evaluation team	6 weeks	Evaluation team
Kick off the evaluation with a stakeholder meeting	2 weeks	Evaluation team, funding partners, national authorities
Gain access to datasets	2 weeks	Evaluation team
Conduct preliminary analysis	3 weeks	Evaluation team
Develop report outline	1 week	Evaluation team
Complete rest of analysis plan	4 weeks	Evaluation team
Develop draft report and share with core stakeholders	4 weeks	Evaluation team
Allow core stakeholders to review draft report	4 weeks	Core stakeholders
Convene consultative meeting to present preliminary results	1 week	Core stakeholders, evaluation team
Develop second draft report, incorporating feedback	4 weeks	Evaluation team
Allow external reviewers to comment on report	4 weeks	Selected stakeholders
Complete final edits, proofreading, and formatting	4 weeks	Evaluation team, editor, proofreader, graphic designer
Print report	1 week	Printer
Hold launch event to share findings	2 weeks	Evaluation team, national authorities

III.3.2 Human Resources and Skills

The staff for the evaluation can come from the NMCP, international agencies, academic institutions, or other malaria control partners. An implementing partner or consultants can be brought in to perform a specific task, and an in-country implementing partner can help coordinate the evaluation. Ideally, an in-country implementing partner will know how to help the NMCP hold stakeholder meetings, access in-country datasets, analyze data, and organize report reviews in country. Individuals from the local implementing partner can aid the NMCP in collecting background information on the country and its history of malaria control.

The evaluation team should have at least one person with epidemiology and biostatistics skills, preferably someone who understands malaria control. This person should know how to work with a key statistical package such as STATA or SAS, and how to analyze large population-based household survey data. The analyst will work throughout the evaluation process, with at least a 50% level of effort during several months. The evaluation team also should consider finding someone to help manage and clean data. This person would work early in the evaluation, with a 50% level of effort during data compilation and cleaning stages. Depending on the types of data available and the analysis chosen, it also may be useful to have a member of the evaluation team who is a geographic information system (GIS) expert for select analyses (~10% level of effort). The analyst also can write the report, or the team can hire additional staff to write the report.

III.3.3 Financial Resources

Costs associated with an evaluation include staff, stakeholders meetings, data access, printing, and dissemination. Most costs for conducting an evaluation come from hiring staff to complete the analysis and report writing, coordinate the evaluation in country, reserve a meeting venue, or prepare materials for stakeholders meetings. Optimally, datasets are available publically or are available to evaluation partners, but it may be necessary to pay for some datasets. Costs associated with finalization of the report depend on the format of the end product (electronic or hardcopy). If a dissemination event to announce the main findings and availability of the report is a possibility, the team should factor it into the budget from the outset.

---

[1] Bamberger M, Rugh J, and Maybry, L. Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints. 2nd edition. 2012. SAGE Publications, Inc.

[2] Tamburrini A, Gilhuly K, and Harris-Roxas B, Enhancing benefits in health impact assessment through stakeholder consultation. Impact Assessment and Project Appraisal, 2011. 29(3):195–204.

[3] United Nations Development Programme. Handbook on Planning, Monitoring and Evaluating for Development Results. 2009.

[4] Bryson JM. What to do when stakeholders matter: Stakeholder identification and analysis techniques. Public Management Review. 2004. 6(1):21–53. DOI: 10.1080/14719030410001675722.

[5] Rowe AK, et al. Viewpoint: Evaluating the impact of malaria control efforts on mortality in sub-Saharan Africa. Trop Med Int Health. 2007. 12(12):1524–1539, doi:TMI1961 [pii] 10.1111/j.1365–3156.2007.01961.x.

[6] Bryce J, Victora CG, Habicht JP, Vaughan JP, and Black RE. The multi-country evaluation of the integrated management of childhood illness strategy: Lessons for the evaluation of public health interventions. *Am J Public Health*. 2004. 94(3): 406–415.

[7] Victora CG et al. Context matters: Interpreting impact findings in child survival evaluations. *Health Policy Plan*. 2005. 20(Suppl 1):i18–i31. doi:20/suppl_1/i18 [pii] 10.1093/heapol/czi050.

[8] Rowe AK, Onikpo F, Lama M, Osterholt DM, and Deming MS. Impact of a malaria-control project in Benin that included the integrated management of childhood illness strategy. *Am J Public Health*. 2011. 101(12):2333–41. doi:AJPH.2010.300068 [pii] 10.2105/AJPH.2010.300068.

[9] Victora CG, Black RE, Boerma JT, and Bryce J. Measuring impact in the Millennium Development Goal era and beyond: A new approach to large-scale effectiveness evaluations. *Lancet*. 2011. 377(9759):85–95. doi:S0140–6736(10)60810–0 [pii] 10.1016/S0140–6736(10)60810–0.

[10] Centers for Disease Control and Prevention. Framework for program evaluation in public health. 1999. MMWR Recomm Rep. 17;48(RR–11):1–40.

IV. Evaluation Design

Key Points

- It is challenging to attribute changes in all-cause child mortality to malaria interventions due to lack of malaria-specific mortality.
- Experimental design in the context of national-scale interventions may not be feasible because it would be difficult to define a control group.
- The overall impact evaluation should rely on an ecological study design, often referred to as a plausibility study design, which assesses simultaneous changes in intervention coverage and malaria burden.

IV.1 Aim of the Impact Evaluation in Each Country

Impact evaluations have several aims, but these are primary objectives:

1. Measure the extent to which malaria control interventions have been implemented and scaled up, as measured against targets set in the national strategic plan and international goals.
2. Assess malaria-related morbidity and mortality before, during, and after scale-up of malaria control interventions.
3. Assess the plausible attribution of the scale-up of malaria control interventions to changes in malaria-related morbidity, and all-cause and malaria-related child mortality.

IV.2 Constraints in Evaluating Malaria Programs

Most malaria control programs are full coverage and intend to reach all populations at risk for malaria. This means a contemporaneous control group is not available for use in evaluations; areas without malaria control scale-up, but with characteristics similar to the scale-up areas, simply do not exist. Without a control or comparison group, it is difficult to assess what would have happened if the national program had never been scaled up. Thus, in most circumstances, it is hard to infer direct causation between exposure to malaria control interventions and observed changes in the malaria burden.

In most circumstances, it is impossible to quantify the relative contribution of different parts of national malaria control programs—such as insecticide-treated nets (ITNs) compared with indoor residual spraying (IRS)—to any observed reduction in malaria burden, because most national programs have a package of malaria control interventions, either simultaneous or on a staggered schedule.

The evaluation will describe the contributions to the malaria control program of donors, such as the Global Fund to Fight AIDS, Tuberculosis and Malaria compared with the President's Malaria

Initiative (PMI). In most circumstances, however, it is impossible to quantify the relative contribution of different donor funds and program implementers in a country to any observed reduction in malaria burden.

All existing data collection systems have limitations. Therefore, MERG recommends triangulating multiple data sources and impact-related indicators rather than using only one data source. Even so, it may not be possible to determine if malaria control interventions are reducing all-cause child mortality (ACCM) or severe anemia in children.

IV.3 Study Design Options

Table IV.1 summarizes the most common evaluation study design options available for national public health programs, including causal inference obtained, feasibility, strengths, and weaknesses. An experimental evaluation design (individual- and community-randomized control trials) is the most rigorous evaluation design to assess a causal relationship between malaria program exposure and changes in the malaria burden, with the use of random assignment for areas that receive and do not receive malaria control interventions. This evaluation design is followed in rigor by a quasi-experimental evaluation design, which may use an independent comparison group or the intervention group as its own comparison group.[1] Data from a comparison group in these designs serve as a counterfactual, which allows an evaluator to assess what would have happened if the intervention never occurred, typically using a difference-in-difference analysis.[2,3] A causal relation would then be defined by showing that any observed decreases in the malaria burden in program areas directly resulted from malaria control interventions, not from extraneous factors.

Reality, however, prevents use of a control or comparison group in national intervention evaluations. Because most malaria control programs were scaled up as full-coverage programs among those at risk of malaria, randomized control trials are nearly impossible. It would be unethical to withhold proven, effective interventions from certain members of the population. In cases where malaria control programs were scaled up in a staggered fashion, which might allow for the use of an experimental or quasi-experimental design, rollout was nearly always done first in areas with the highest transmission, preventing any comparison between those with and without the interventions.

Although statistically constructed controls through matching and regression discontinuity can help in assessing causality and are promoted for use in these evaluations, these designs are limited to a subpopulation in countries with such available data. In some circumstances, a dose-response design might be feasible where malaria control interventions resulted in varying levels of subnational coverage. This is perhaps the strongest feasible design for use at a national level. However, the lack of continuous, valid, population-level data on malaria program intensity and malaria burden at subnational levels renders this design difficult to use in most cases. Both constructed comparison and dose-response evaluation designs are excellent methods to bolster less rigorous designs; however, they do not work for stand-alone, full-country evaluations across most countries.

TABLE IV.1: SUMMARY OF STUDY DESIGN OPTIONS TO EVALUATE COUNTRY MALARIA CONTROL PROGRAMS

Study Design		Causal Inference	Feasibility	Strengths	Weaknesses	Key Assumptions
Experimental	Community randomized control trial (CRCT): pretest/posttest, posttest only designs, and interrupted time series with randomized control	High (Strongest)	Low	High internal validity Strong evidence of impact Relatively simple analysis (difference-in-differences)	• Does not lend itself to full-coverage programs, or where other programs are rolling out • Unethical to intentionally withhold proven interventions, so often impossible to use	• Randomization produces equivalent groups • External influences affect both groups equally • Same intensity of treatment • Assignment to experimental groups does not itself alter the behavior of service providers or study subjects • Limited contamination
Quasi-experimental	Nonequivalent control/ nonrandomized step-wedge and interrupted time series with nonrandomized control	High	Medium	Useful for phased program rollout Strong evidence of impact High internal validity Allows conduct of difference-in-difference analysis to assess impact	• Selected areas for initial rollout may differ on key characteristics to areas in subsequent phases (selection bias) • Unethical to intentionally withhold proven interventions, so not possible to use in many situations • Does not lend itself to full-coverage programs or where other programs are rolling out	• Project influences when and where the phased rollout starts • Confounding factors can be measured and controlled for in analysis
	Constructed control: matching and discontinuity designs, instrumental variables	Medium	Medium	Attempts to account for selection bias through statistical analysis Versatility Some evidence of impact Medium internal validity	• Vulnerable to selection bias • Advanced statistical analyses required • Local average treatment effect cannot always be generalized	• Confounding factors can be measured and controlled for in analysis • Constructed control yields unbiased treatment effect • Constructed control has produced equivalent groups • External influences affect both groups equally • Same intensity of treatment • Extraneous factors do not differentially alter the behavior of service providers or study subjects

Study Design		Causal Inference	Feasibility	Strengths	Weaknesses	Key Assumptions
Observational	Dose response	Medium-low	Medium-high	• Versatile: estimates of impact derived at multiple levels • Modest internal validity	• Selection bias • No counterfactual • Differential participation bias • Differential attrition bias	• Confounding factors can be measured and controlled • Exposure to treatment will vary
	Interrupted time series without a control group, reflexive control: time series and repeated measures designs (plausibility design)	Low-medium	High	• Treatment group serves as its own control over time; easy to implement • No need to exclude group from treatment • Versatility • Useful for full coverage programs	• No counterfactual, so cause and effect cannot be inferred • Must measure and account for confounding factors in analysis to establish plausible impact • Unobserved heterogeneity • Differential selection bias and attrition	• Program is preexisting or full coverage • Pretest measures are valid estimates of the counterfactual • Treatment effect demonstrated if posttest measures significantly differ from pretest measures • No other plausible explanations for observed treatment effects

IV.4 Plausibility Design

To meet the evaluation objectives with the constraints outlined above, the overall evaluation should rely on a pre-post, reflexive-control study design—often referred to as a plausibility study design—which typically consists of a nonexperimental approach based on a pretest and posttest with a plausibility argument.[4,5,6] Such a design aims to assess simultaneous changes in intervention coverage and malaria burden at the population level. An important aspect of this design for evaluating the effects of the scale-up of a malaria control program is that the interventions being assessed all have been empirically proven to have an impact on the outcomes being measured; for example, ITNs and their established protective efficacy on reducing ACCM from randomized controlled trials.[7]

However, efficacy may not hold in some contexts, such as insecticide resistance. In these cases, MERG recommends properly documenting the change in efficacy. If intervention coverage is shown to increase sufficiently with simultaneous improvements in defined morbidity and mortality outcomes—both measured at the population level—then it would be plausible that the program contributed to improved outcomes, assuming contextual factors can be accounted for and the timing of the intervention scale-up matches changes in population health outcomes (ACCM measured by household surveys reflects a period on average of 2.5 years preceding the survey). The counterfactual is that changes in outcomes would not have happened, at least to the extent observed, without scaling up the malaria control program. This conclusion is strengthened if the

evaluation team establishes baseline measures of malaria burden outcomes before scale-up of the malaria control program.

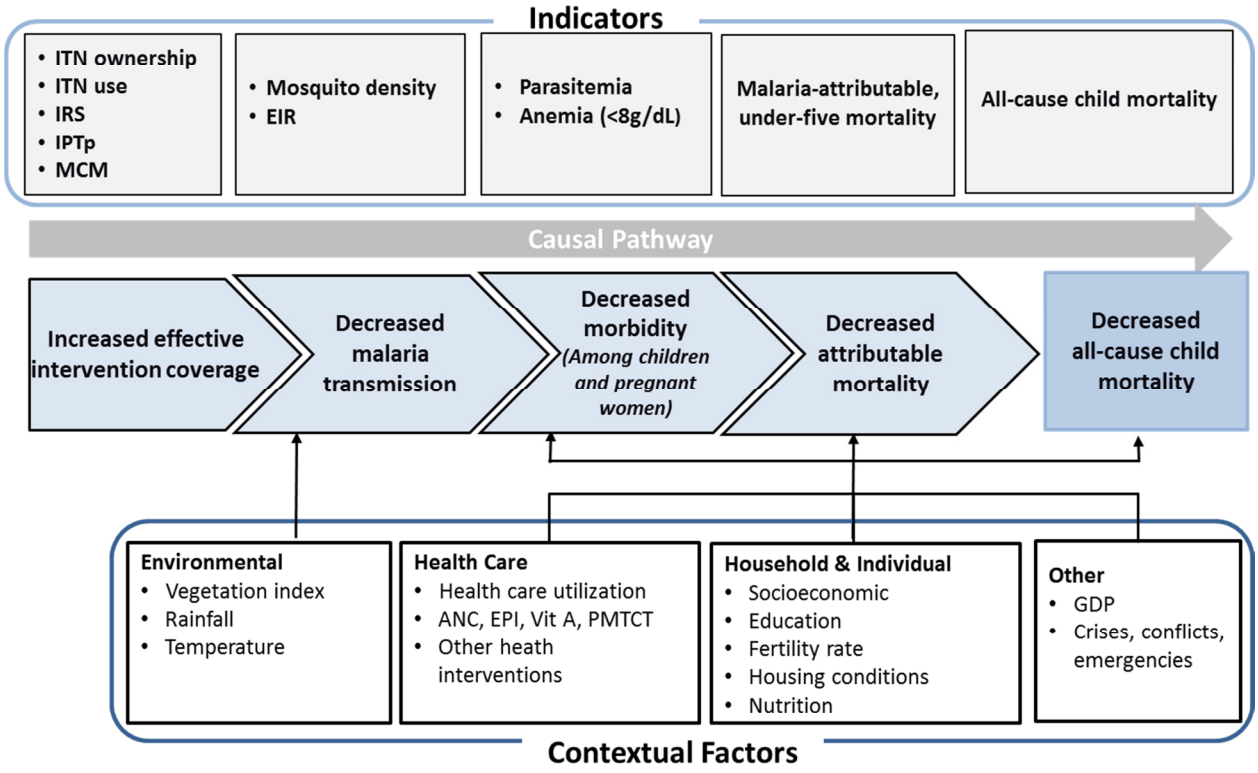
Consistent, equitable patterns of intervention coverage and improvements in primary impact indicators across multiple data collection systems while measuring and accounting for potential contextual factors strengthen the plausibility study design. Details of synthesizing data and secondary data analysis to support the plausibility study design appear in Section XI. All available data points of reasonable validity from the formal impact evaluation are considered in establishing the overall direction and magnitude of the impact of the malaria control program on key impact indicators. Doing so establishes the level of evidence for causality based on the framework established by Bradford Hill[8] and interpreted by Hofler[9] to systematically incorporate all available data points of reasonable validity into a comprehensive image of the plausible impact the malaria control program has on the chosen impact indicators.

Many limitations and constraints hinder the use of a plausibility study design. Appendix A provides details, summarized here, of the major limitations and their associated assumptions and potential means of mitigation, where possible:

- Preclusion of causal inference from the plausibility study design
- Translation of efficacy to effectiveness under program conditions
- Difficulty in measuring external contextual and confounding factors
- Lack of sufficient data points to draw meaningful conclusions

The conceptual framework in Figure IV.1 shows the relationships between empirically proven malaria control interventions (top)[7,10,11,12,13,14,15] and external factors (bottom) on the desired impacts of reductions in morbidity and mortality resulting from malaria and all causes. It is critical to monitor such external, potentially contextual, factors to carefully interpret changes in malaria morbidity and ACCM. Other than malaria intervention coverage, rainfall, urbanization, and temperature are the primary contextual factors that influence malaria transmission, and thus, morbidity and child mortality resulting from malaria and all causes. The evaluation must consider many key external factors that influence ACCM, including population coverage of oral rehydration therapy (ORT) and zinc for diarrheal disease; antibiotics for pneumonia; Expanded Programme on Immunization (EPI); prevention of mother-to-child transmission of HIV (PMTCT); nutrition interventions; access and use of maternal, newborn, and child health services; and improvements in socioeconomic status of the household and community. Further discussion of contextual factors appears in Section X.

FIGURE IV.1: CONCEPTUAL FRAMEWORK OF KEY FACTORS IN ASSESSING THE IMPACT OF MALARIA CONTROL PROGRAMS ON MALARIA AND ALL-CAUSE CHILD MORTALITY, USING A PLAUSIBILITY STUDY DESIGN



Notes: ANC=antenatal care, EIR=entomological inoculation rate, EPI=Expended Program on Immunization, ITN=insecticide-treated net, IRS=indoor residual spraying, IPTp=intermittent preventive treatment for pregnant women, GDP=gross domestic product, MCM=malaria case management, Vit A=vitamin A, PMTCT= prevention of mother-to-child transmission of HIV

IV.5 What Results Might Look Like

Results from the Bioko Island evaluation, shown in Appendix F,[16] show perhaps the most consistent and impressive results expected from a plausibility evaluation study design using multiple data points after successful, full-scale scale-up of a national malaria control program. In this example, which used a single-group pre-post design with plausibility argument to evaluate four years of high-coverage scale-up of IRS, ITNs, and improved case management on the island, results were consistent and in expected directions. The evaluation showed simultaneous drops in malarial parasite infection, anemia, and fever history prevalence in children following the scale-up. The evaluation team also observed coinciding declines in ACCM that matched temporally with the 2.5-year lag period that resulted from the use of birth histories from household surveys. Similarly consistent results also were observed in Zanzibar, as shown in Appendix H.[17]

This consistency is an unlikely scenario that should not be expected in most evaluations that use a plausibility study design; in reality, simultaneous changes in all key health impact indicators likely will

not all move in expected directions. For example, in Zambia, malaria parasite prevalence declined from a national mean of 22% to 12% from 2006 to 2008 following national scale-up of ITNs.[18] In 2010, however, the national prevalence went up to 16%. At first, it appears the program failed; however, the proper analysis demonstrated that if ITNs had not been in place, the malaria parasite prevalence in 2010 likely would have been even worse. This analysis accounted for lower-than-normal rainfall in 2008, higher-than-normal rainfall in 2010, and the observation that some districts in Zambia had ITNs older than three years on average.

Another scenario that could occur is declining ACCM that preceded the scale-up of malaria control programs. This situation occurred all over Africa, where ACCM has been falling since the 1980s. This doesn't necessarily mean malaria control programs will be shown to have no impact on ACCM, but evaluators should carefully match the timing of measures of ACCM with program implementation. The counterfactual is a declining trend in ACCM from baseline, but at a lower rate than would have occurred with scale of malaria intervention.

Also in Zambia, trends in laboratory-confirmed incidence of outpatient malaria cases appear to have increased since 2009. However, simultaneous increases in access to health services, scale-up of rapid diagnostic tests nationwide, and improvements in health management information systems can easily explain the trends. These are but a few examples of the complexity of using a plausibility design to assess the impact of full-coverage, national malaria control programs. Results likely will be messy and require rigorous analysis of multiple data points at once, accounting for important contextual and potentially confounding factors. Several options for such analyses appear in Appendix D.

---

[1] Fretheim A et al. Interrupted time-series analysis yielded an effect estimate concordant with the cluster-randomized controlled trial result. *J Clin Epidemiol*. 2013. 66(8):883–7.

[2] Khandker RS, Koolwal GB, Samad HA. *Handbook on Impact Evaluation: Quantitative Methods and Practices (World Bank Training Series)*. 2010. The International Bank for Reconstruction and Development / The World Bank, Washington DC. eISBN: 978-0-8213-8029-1, DOI: 10.1596/978-0-8213-8028-4.

[3] Gertler PJ, Martinez S, Premand P, Rawlings LB, and Vermeersch CMJ. *Impact Evaluation in Practice*. 2011. The International Bank for Reconstruction and Development / The World Bank, Washington DC. eISBN: 978-0-8213-8593-7, DOI: 10.1596/978-0-8213-8541-8.

[4] Rowe AK, Steketee RW, Arnold F, Wardlaw T, Basu S, Bakyaita N, et al. Viewpoint: Evaluating the impact of malaria control efforts on mortality in sub-Saharan Africa. *Trop Med Int Health*. 2007. 12(12):1524–1539.

[5] Habicht JP, Victora CG, and Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol*. 1999. 28(1):10–18.

[6] Victora CG et al. Measuring impact in the Millennium Development Goal era and beyond: A new approach to large-scale effectiveness evaluations. *Lancet*. 2011. 377(9759):85–95.

[7] Lengeler C. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev*. 2004. (2):CD000363.

[8] Hill AB. The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*. 1965. 58(5): 295–300.

[9] Hofler M. The Bradford Hill considerations on causality: A counterfactual perspective. *Emerging Themes in Epidemiology*. 2005. 2:11.

[10] Lindblade KA, et al., 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*. 2004. 291(21):2571–2580.

[11] ter Kuile FO, et al., Impact of permethrin-treated bed nets on malaria and all-cause morbidity in young children in an area of intense perennial malaria transmission in western Kenya: Cross-sectional survey. *Am J Trop Med Hyg*. 2003. 68(4 Suppl):100–107.

[12] Kidane G, and Morrow RH. Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: A randomised trial. *Lancet*. 2000. 356(9229):550–555.

[13] ter Kuile, FO, van Eijk AM, and Filler SJ. Effect of sulfadoxine-pyrimethamine resistance on the efficacy of intermittent preventive therapy for malaria control during pregnancy: A systematic review. *JAMA*. 2007. 297(23):2603–2616.

[14] Gamble CL, Ekwaru JP, and ter Kuile FO. Insecticide-treated nets for preventing malaria in pregnancy (Review). *Cochrane Database Syst Rev*. 2009. 2:1–51.

[15] Pluess B, et al. Indoor residual spraying for preventing malaria. *Cochrane Database Syst Rev*. 2010. 4:CD006657.

[16] Kleinschmidt I, Schwabe C, et al. Marked increase in child survival after four years of intensive malaria control. *Am J Trop Med Hyg*. 2009. 80(6):882–888.

[17] Bhattarai A, Ali AS, et al. Impact of artemisinin-based combination therapy and insecticide-treated nets on malaria burden in Zanzibar. *PLoS Med*. 2007. 4(11):e309.

[18] Chizema-Kawesha E, Miller JM, et al. Scaling up malaria control in Zambia: Progress and impact 2005–2008. *Am J Trop Med Hyg*. 2010. 83(3):480–488.

V. Program Description

Key Points

- The malaria control program description puts the evaluation into context and provides information to interpret evaluation findings.
- The description of the malaria control program should include basic country information, a description of the health care system, a malaria epidemiologic profile of the country, a timeline for implementing malaria control interventions and policies, funding and commodity inputs, and key events in the country.

As mentioned in Section III.1.2, a malaria control program description is developed as part of the early steps of conducting an impact evaluation. It also provides an important section in the evaluation report. A malaria control program description should cover the following topics: basic country information, description of the health care system, a malaria epidemiologic profile of the country, a timeline for implementing malaria control interventions and policies (strategies), funding and commodity inputs, and key events in the country. The narrative should also note the timing of major malaria surveys.

V.1 Country Background

The country background section should provide a map and description of administrative boundaries in the country. It is important to show boundaries at a level relevant to the analysis, such as a regional, provincial, zonal, or district level. The background section also should provide information about the geography and seasons, which will be important for interpreting survey results. The main source of income in the country and some general development indicators, such as United Nations Development Programme's (UNDP) Human Development Index and Gross Domestic Product (GDP), should be included to provide some context for the country.

V.2 Health Care System

This section of the description explains the structure of the health care system, with information on the number of facilities at each level of the system and how this has changed over time, particularly over the evaluation period. A description of policy changes, such as free access and insurance programs, as well as when and how the policy was implemented and scaled up, should be included. It is important to understand the proportion of the population that seeks care from the public sector compared with the private sector, because this can guide the interpretation of case management results. This section should also cover the number and ratio (health worker: population) of each type of health care worker, including the roles of community health workers in countries with community

case management. Information on how the health care system has evolved over the evaluation period deserves attention as well.

V.3 Malaria Epidemiologic Profile

The background chapter of a malaria impact evaluation should have a malaria situation analysis, which includes a breakdown of the country into malaria risk zones. For analysis stratified by malaria risk (see Section VIII), it helps to know if there are malaria risk zones that the National Malarial Control Program (NMCP) uses. For example, in Angola, the Malaria Indicator Survey (MIS) is sampled according to four risk zones: hyperendemic, mesoendemic stable, mesoendemic unstable, and Luanda (the capital city). A map depicting the malaria risk zones should appear in this section of the report, especially if the map shows changes in the malaria epidemiologic profile over the evaluation period; however, these maps are not often available for many points in time.

The malaria epidemiologic profile for a country should give a general sense of the numbers of malaria cases and all-cause or malaria-specific deaths (all ages and children under age 5 years). The profile should distinguish between laboratory-confirmed malaria cases and clinical cases. This is important because some countries might not have set the policy for confirmation in the early years of the evaluation period. The morbidity and mortality analysis sections of an evaluation report can provide a more detailed analysis of trends in malaria cases and deaths.

Another part of the description covers the species of malaria parasites and mosquito vectors found in the country.

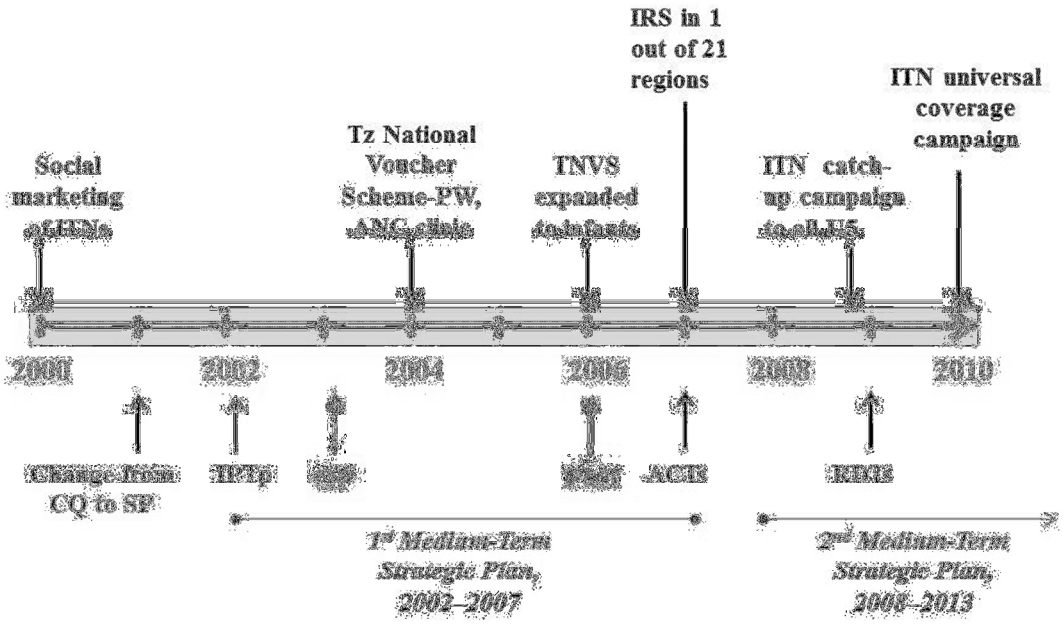
V.4 Malaria Control Strategy

The background should describe the history of malaria control in the country, beginning with a description of early efforts and policies, if they existed, for malaria control and then a description of how the NMCP was formed. Often a country has an advisory body that supports the NMCP, and this body should be described, along with a mention of malaria research or operational research projects conducted in the country that inform malaria policy decisions.

The malaria control strategy section of the evaluation report should highlight national strategic plans, including goals, targets, and activities. The section also should give an overall sense of milestones of malaria control in the country over the evaluation period. Figure V.1 shows an example of how to present this information. In the figure, introduction of vector control measures insecticide-treated nets (ITNs) and indoor residual spraying (IRS) are shown across the top of the timeline. Changes in drug policy and introduction of intermittent preventive treatment for pregnant women (IPTp), artemisinin combination therapies (ACTs), and rapid diagnostic tests (RDT) are shown below the timeline. If a specific intervention is rolled out in only certain areas, this geographic information should be mentioned. Input of funding from two of the donors, the Global Fund to Fight AIDS, Tuberculosis and Malaria (Global Fund) and President's Malaria Initiative

(PMI), is shown in red below the timeline. During this evaluation period, Tanzania developed two malaria strategic plans that guided malaria control interventions in the country. Not shown in Figure V.1 are the national surveys and other data sources, which also should be included in a milestone or timeline figure.

FIGURE V.1: MILESTONES IN MALARIA CONTROL, MAINLAND TANZANIA, 2000–2010 (EXAMPLE)



Notes: Insecticide-treated net (ITN), Tanzania (Tz), pregnant women (PW), antenatal care (ANC), Tanzania National Voucher Scheme (TNVS), indoor residual spraying (IRS), children under five years of age (U5), chloroquine (CQ), sulfadoxine/pyrimethamine (SP), intermittent preventive treatment in pregnancy (IPTp), The Global Fund to Fight AIDS, Tuberculosis, and Malaria (Global Fund), President's Malaria Initiative (PMI), artemisinin combination therapy (ACT), rapid diagnostic test (RDT)

Source: Evaluation of the Impact of Malaria Control Interventions on All-Cause Mortality in Children Under-5 in Mainland Tanzania. Tanzania Malaria Impact Evaluation Group. Supporting evaluation documents, 2012.

A description of intervention implementation is necessary in a malaria impact evaluation. It can appear in the malaria control strategy section of the evaluation report or as part of the section that measures coverage of interventions. It should include implementation strategies, such as ITN implementation through mass distribution campaigns, routine distribution through Expanded Programme on Immunization (EPI) and antenatal care visits, social marketing, and retail and policy changes, such as first-line antimalarial drugs and shifts in IRS insecticide.

V.5 Financial Resources

The malaria control program section of the evaluation report should describe funding sources for malaria control. The government's financing of malaria control should be documented, and all external donor funding should be accounted for. Ideally, disbursements by year for each donor agency will be available. The NMCP and donor agencies provide government and donor funding

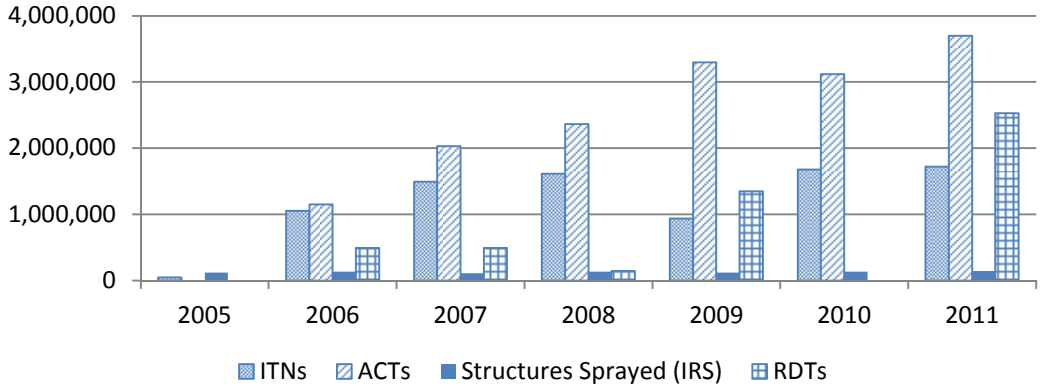
information. This section should document malaria health expenditures when available. This information may be available from National Health Accounts. The report should include such indicators as these, for each year available:

- Total expenditure on health for malaria
- Total expenditure on malaria as a percentage of total health expenditure
- Total government expenditure on health for malaria
- Per-capita total expenditure on health for malaria
- Total expenditure on health for malaria as a percentage of GDP
- Government expenditure on health for malaria as a percentage of GDP
- Government per-capita total health expenditure
- Government total expenditure on health as a percentage of total government expenditure
- Per-capita national expenditure on health, converted to an average U.S. dollar exchange rate

V.6 Commodity Inputs

The evaluation report must document the country's commodity inputs, such as ITNs, ACTs, and RDTs, during the evaluation period. Make note of whether commodity numbers refer to procured or distributed commodities (distributed preferred). The document can include the number of structures sprayed during IRS, though IRS is not strictly expressed as a commodity. This information is available from the NMCP and donor agencies. An example of commodities distributed or procured is shown in Figure V.2.

FIGURE V.2: MALARIA COMMODITIES PROCURED OR DISTRIBUTED IN ANGOLA, 2005–2011 (EXAMPLE)



Notes: Insecticide-treated nets refer to the number distributed; artemisinin combination therapies and rapid diagnostic tests refer to the number procured; and indoor residual spraying refers to the number of structures sprayed.

Source: RBM Progress & Impact Series: Focus on Angola. Draft, May 2013.

V.7 Key Events

The evaluation report also should include a section on key events in the country to provide context for malaria control program implementation. This section of the report should include political factors, such as civil wars and migration; environmental and climate events, such as floods, droughts, and natural disasters; and major disease outbreaks.

VI. Measuring the Coverage of Malaria Control Interventions

Key Points

- Large population-based surveys, such as the Demographic and Health Surveys (DHS), Malaria Indicator Surveys (MIS), and Multiple Indicator Cluster Surveys (MICS), are key sources of coverage data. Other sources, including routine health information systems, also can provide coverage data or other complementary information.
- Here are some aspects to consider when measuring the coverage of malaria control:
  1. Survey timing: DHS and MICS data generally are collected during the dry season, and MIS is conducted during peak transmission season; consider the timing when comparing indicators from different types of surveys.
  2. Malaria endemicity: Evaluators should stratify their analyses by intervention coverage in the highest-risk populations and in-country variation of endemicity.
- Routine data offers information about service delivery and facility performance, which helps measure coverage.

Roll Back Malaria's table of outcome indicators (Table VI.1) gives an overview of the updated objectives, targets, milestones, and coverage indicators. These can guide implementation of malaria control programs and form a framework for monitoring and evaluation. Coverage indicators are generally measured using population-based data (in bold). However, routine health data might also contribute to understanding intervention coverage.

TABLE VI.1: RBM OUTCOME INDICATORS TO MONITOR MALARIA CONTROL INTERVENTIONS

Key Indicator	Further Analysis	Supporting Indicator
→ Inpatient malaria deaths per 1000 persons per year	→ Has health facility reporting completeness changed over time?	→ Completeness of monthly health facility reports
→ **All-cause under 5 mortality rate**	→ What factors are responsible?	→ **Programme coverage indicators in this table (detailed below)**
→ Proportion of suspected malaria cases that receive a parasitological test		
→ **Proportion of children under 5 years old with fever in the last 2 weeks who had a finger or heel stick**	→ Are people seeking advice or treatment for fever and from where?	→ **Proportion of children under 5 years old with fever in the last 2 weeks for whom advice or treatment was sought**
→ Proportion of confirmed malaria cases that receive first-line antimalarial treatment according to national policy	→ Are adequate quantities of antimalarial medicines available?	→ Proportion of health facilities without stock-outs of key commodities by month
→ **Proportion receiving first-line treatment among children under 5 years old with fever in the last 2 weeks who received any antimalarial drugs**		
→ Confirmed malaria cases (microscopy or RDT) per 1000 persons per year	→ Has diagnostic effort changed over time?	→ Annual blood examination rate
	→ Has health facility reporting completeness changed over time?	→ Completeness of monthly health facility reports
	→ Have test positivity rates changed over time?	→ Malaria test positivity rate
→ **Parasite prevalence: proportion of children aged 6–59 months with malaria infection**	→ Is there other evidence of morbidity change?	→ **Proportion of children aged 6–59 months with a hemoglobin measurement of <8 g/dL**
→ **Proportion of population with access to an ITN within their household**	→ How many households have at least one ITN?	→ **Proportion of households with at least one ITN**
	→ How many households have enough ITNs for each occupant?	→ **Proportion of households with at least one ITN for every two people**
	→ Were enough ITNs delivered to ensure at least one ITN per two people at risk?	→ Proportion of population at risk potentially covered by ITNs distributed
	→ Are specific risk groups receiving ITNs?	→ Proportion of targeted risk group receiving ITNs
→ **Proportion of population that slept under an ITN the previous night**	→ Are specific population groups using ITNs?	→ **Proportion of children under 5 years old who slept under an ITN the previous night**
		→ **Proportion of pregnant women who slept under an ITN the previous night**
	→ Are available ITNs being used?	→ **Proportion of existing ITNs used the previous night**
→ Proportion of population protected by IRS within the last 12 months		
→ **Proportion of households with at least one ITN for every two people and/or sprayed by IRS within the last 12 months**	→ How many households have been reached with at least one vector control method?	→ **Proportion of households with at least one ITN and/or sprayed by IRS within the last 12 months**
→ **Proportion of women who received 3 or more doses of intermittent preventive treatment for malaria during ANC visits during their last pregnancy*** **	→ Is IPTp received by all pregnant women at each scheduled ANC visit?	→ Proportion of women attending ANC who received 1, 2, 3, or 4 doses of IPT***
→ Percent of districts reporting monthly numbers of suspected malaria cases, number of cases receiving a diagnostic test and number of confirmed malaria cases		
→ Number of new countries in which malaria has been eliminated	→ What are the trends in malaria cases?	→ Number of active foci reported per year
		→ Number of cases by classification (indigenous, introduced, imported, induced)
	→ How strong are surveillance systems?	→ Proportion of private facilities reporting to national malaria surveillance system

■ Indicator derived from household surveys

Source: World Malaria Report. Geneva, World Health Organization, 2012.

VI.1 Outcome Indicators Derived from Population-based Household Surveys

The Roll Back Malaria Monitoring and Evaluation Reference Group (RBM MERG) provides guidance on outcome indicators that can be tracked by countries through population-based household surveys in the *Household Survey Indicators for Malaria Control,*[1] which is updated regularly. New guidelines recommend the use of 13 primary outcome indicators to monitor the scale-up of key malaria control interventions, such as vector control using insecticide-treated nets (ITNs) and indoor residual spraying (IRS), prompt and effective treatment and use of diagnostics, and prevention and control of malaria in pregnant women (Table VI.1, indicators in bold). Other indicators might be suggested for a specific evaluation or for analyses in a particular country. The guidelines include supplemental indicators that also may be suitable at the country level.

Primary sources of intervention coverage data for impact evaluations consist of large population-based 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 three to five years and provide national, and sometimes subnational, estimates for key malaria control indicators. Other surveys and data sources, such as post-campaign, long-lasting insecticidal nets (LLINs) coverage surveys, demographic surveillance sites, and routine health information systems, might also contribute to understanding intervention coverage; data from these sources should be evaluated and incorporated in the analysis when possible.

Analysts can calculate coverage for each of the interventions for each survey period directly from survey datasets, or they can extract information from published reports of surveys when datasets are not available. When possible, each estimate should have 95% confidence intervals. Analysts can also use standard statistical methods to assess changes in proportions over time and determine changes in intervention coverage. If data are available for more than two time points, a trend analysis can be performed. Analysis of intervention coverage also should take into account malaria transmission zone, urban and rural areas, and wealth quintiles.

Criteria for including data from non-DHS, MIS, and MICS surveys in the analysis should include the following: (1) survey methodology available for review to determine sampling scheme and representativeness, (2) availability of data stratified or age information to permit extracting results for children under age 5 years, (3) availability of survey sampling weights for secondary analysis of existing data, and (4) no more than 5% of missing data for indicators of interest.

VI.1.1 Considerations

Analysis should consider the following aspects:

- **Survey timing.** DHS and MICS usually are conducted during the dry season, and therefore, outside of the peak malaria transmission season. MIS usually are conducted during the peak

malaria transmission season. Because of these differences in survey timing, evaluators should take care when comparing indicator measurements from different types of surveys or data sources. For example, ITN use might be higher during the peak malaria transmission season than during the dry season; therefore, a MIS survey could yield a higher estimate of ITN usage than a DHS survey.

- **Malaria endemicity.** In some countries, only part of the population is at risk of malaria; thus, national surveys might underestimate the true effective coverage among populations at risk. An appropriately stratified analysis allows evaluators to look at intervention coverage by malaria transmission zone and in the highest-risk populations, such as by wealth quintiles or urban and rural residences. Levels before scale-up of malaria control interventions should form the basis of these stratifications, for use throughout the evaluation period.
- **Measurement of vector control intervention coverage.** Six vector control indicators, gathered in population-based surveys, examine ownership and use of ITNs and coverage of IRS in high-risk groups (children under age 5 years and pregnant women) and by the entire population.
- **Prompt, effective treatment and use of diagnostics.** Previous versions of the core RBM treatment indicators in the *Household Survey Indicators for Malaria Control* used fever as a proxy measure for malaria. These indicators, which were used widely in the past but are no longer recommended, are (1) proportion of children under age 5 years with fever in the past 2 weeks who received any antimalarial treatment, and (2) proportion of children under age 5 years with fever in the past 2 weeks who received antimalarial treatment according to national policy within 24 hours from onset of fever. As the use of diagnostics has scaled up, these indicators have become difficult to interpret because parasitological confirmation can determine that some fever cases are not malaria. These indicators are now replaced with indicators on care seeking and first-line antimalarial treatment among children under age 5 years with fever who received any antimalarial drugs. Changes in the indicators for prompt and effective treatment and use of diagnostics should be interpreted in the context of what is known about intervention scale-up in each country.

VI.2 Outcome Indicators from Routine Health Information Systems

Data from routine health information systems, either malaria-specific or integrated, may also contribute to understanding intervention coverage. Whereas population-based data can provide evidence of coverage and impact even outside of the public health system, routine data offers information about service delivery and facility performance. This includes information on malaria test positivity, facility reporting, stockouts, and adherence to testing and treatment guidelines.

Routine data are useful because program staff can calculate indicators quickly and regularly, allowing them to easily identify issues and make changes to meet new definitions or reporting requirements.

In addition, these systems are already in place, so the financial and time burden to calculate indicators is low.

WHO's Malaria programme reviews: A manual for reviewing the performance of malaria control and elimination programmes[2] outlines the following coverage indicators that can be derived from routine data:

- Number and rate of suspected malaria cases tested for malaria
- Annual blood examination rate
- Proportion of suspected (clinical) cases of malaria tested for malaria
- Proportion of people treated with ACT within 24 hours
- Proportion of epidemics detected within two weeks and responded to within one week

VI.2.1 Considerations

Analysis should consider the following aspects:

- **Seasonality.** Routine data collection and collation makes data less prone to issues related to seasonality.
- **Malaria endemicity.** Numbers of malaria tests and positivity often appear in routine summary reports in high prevalence countries, but not as consistently in countries where the malaria risk is low or localized.
- **Utilization.** The proportion of the target population utilizing reporting facilities and the completeness of recording influences indicators derived from facility-based data. Low or high utilization may not only reflect population density but may also be an outcome of service quality.
- **Consistent definitions.** For the purpose of assessing trends across time and locations, it is important to define indicators clearly. This includes outlining information about how results were compiled and which types of facilities and tests were included.

VI.2.2 Key Reference Documents

- Household Survey Indicators for Malaria Control: http://www.rollbackmalaria.org/toolbox/tool_HouseholdSurveyIndicatorsForMalariaControl.html
- DHS tools: http://www.measuredhs.com/Topics/Malaria-Corner/
- MICS tools: http://www.childinfo.org/mics5_tools.html
- MIS tool package: http://www.malariasurveys.org/
- Malaria Programme Reviews: A Manual for Reviewing the Performance of Malaria Control and Elimination Programmes: http://www.who.int/entity/malaria/publications/atoz/whomprmalariaprogramperformancemanual.pdf

[1] Roll Back Malaria Partnership Monitoring & Evaluation Reference Group. *Household Survey Indicators for Malaria Control.* June 2013.

[2] World Health Organization. Malaria programme reviews: a manual for reviewing the performance of malaria control and elimination programmes (Trial edition). March 2010.

VII. Measuring Transmission Intensity

Key Points

- Malaria risk can be measured using the entomological inoculation rate, parasite prevalence, or parasite incidence in a number of ways.
- These different measures can be used to compare different geographical settings in a country, compare different periods of time (such as rainy season in contrast to dry season, before and after an intervention), or examine trends over time.
- To understand transmission risk, it is important to have regular data on temperature, which plays a major role in the malaria vector life cycle, and rainfall, which allows for mosquito breeding. Because these factors can vary greatly in a small area, they are particularly useful for subnational analyses.

VII.1 Transmission Intensity

Measuring malaria risk has been a concern since the discovery of the link between the mosquito and the parasite. Macdonald[1] introduced the basic reproductive rate (R_0) concept to describe the sustainability of malaria transmission, which expresses the number of new infections from a single case of malaria without immunity. Malaria transmission is sustained if this number is above one.[1] This concept includes the parasitological and mosquito aspects. Vectorial capacity (VC), introduced later by Garrett-Jones,[2] who removed the parasitological aspect from the basic reproductive rate, expresses the daily expected inoculation of humans per infective case.

The most commonly used measure of malaria transmission intensity under field conditions is the entomological inoculation rate (EIR), defined as the number of infective bites per person per unit of time, often per year.[3] In Eritrea, Shililu, et al. used EIR to show that the risk of exposure to infected mosquitoes was heterogeneous and seasonal, with high inoculation rates during the rainy season, but little or no transmission during the dry season.[4] The researchers concluded that EIR can help quantify levels of exposure in different regions of the country and could be used to evaluate vector control strategies. Although several studies have documented the relationship between EIR and malaria risk,[5,6] some controversies remain. In endemic areas, high EIR may not necessarily lead to high malaria morbidity or mortality because a more exposed population can develop immunity, and thus have less susceptibility to clinical malaria (infection can exist without any sign of illness). Researchers also can measure malaria risk by parasite prevalence or parasite incidence. Where parasite prevalence data are not available throughout the evaluation period, modeling can supplement this information.[7,8]

Researchers can use these measures to compare different geographical settings in a country, compare different periods of time (such as rainy season in contrast to dry season, before and after an intervention), or examine trends over time. Table VII.1 gives an overview of methods to assess malaria risk, and Appendix C provides more details. For most countries, collecting information on the entomological indicators at the national level is challenging; therefore, using information available at the subnational level can be informative. Special studies conducted by research institutions, model predictions, and, in some cases, from the national malaria control program entomological unit can provide information on transmission intensity. The impact evaluation can use transmission intensity to better define a country's malaria risk zone and to assess changes in transmission risk over time and across subnational levels.

TABLE VII.1: MALARIA RISK PARAMETERS [9,10]

Parameter	Definition of index
Human biting rate (HBR)	Bites per person per night by vector population
Vectorial capacity (VC)	Expected human inoculations per infective case per unit of time
Basic reproductive rate (R_0)	Expected new infections per cases without immunity
Entomological inoculation rate (EIR)	Number of infective bites per-person, per-time unit
Parasite prevalence	Percentage of infected people
Incidence of malaria morbidity	Number of new cases during a given period

[1] Macdonald G. Appendix I. Mathematical statement in *The Epidemiology and Control of Malaria*. 1957. pp 201, Oxford University Press; London.

[2] Garrett-Jones C. The human blood index of malaria vectors in relation to epidemiological assessment. *Bull World Health Organ*. 1964. 30:241–61.

[3] Onori E and Grab B. Indicators for the forecasting of malaria epidemics. 1980. *Bull World Health Organ*. 58(1):91–8.

[4] Shililu J, Ghebremeskel T, Mengistu S, Fekadu H, Zerom M, Mbogo C, Githure J, Novak R, Brantly E, and Beier JC. High seasonal variation in entomologic inoculation rates in Eritrea, a semi-arid region of unstable malaria in Africa. *Am J Trop Med Hyg*. 2003. 69:607–13.

[5] Beier JC, Killeen GF, and Githure JI. Short report: entomologic inoculation rates and *Plasmodium falciparum* malaria prevalence in Africa. *Am J Trop Med Hyg*. 1999. 61:109–13

[6] Smith TA, Leuenberger R, and Lengeler C. Child mortality and malaria transmission intensity in Africa. *Trends Parasitol*. 2001. 17(3):145–9.

[7] Malaria Atlas Project, http://www.map.ox.ac.uk/

[8] Gosoniu L, Veta AM, and Vounatsou P. Bayesian geostatistical modeling of Malaria Indicator Survey data in Angola. *PLoS ONE*. 2010. 5(3): e9322.

[9] Ye Y, Sankoh O, Kouyate B, Sauerborn R. *Environmental Factors and Malaria Transmission*. 2008. Ashgate.

[10] Gilles HM, and Warrell DA. *Bruce-Chwatt's Essential Malariology, Third edition*. 1993. Edward Arnold, London.

VIII. Measuring Malaria Morbidity

Key Points

- Malaria-related morbidity should be measured using data from both population-based sources and health facilities.
- For population-level morbidity indicators in areas with high-intensity malaria transmission, measurements should be taken for anemia (hemoglobin value < 8g/dL) and malaria parasite prevalence in children ages 6 to 59 months to assess malaria-related morbidity.
- Health facility-based morbidity indicators help assess malaria-related morbidity, notably in children ages 5 years and under.
- To help interpret data, health facility-based morbidity indicators should be supported by indicators of laboratory testing coverage; specifically, the malaria test positivity rate and proportion of suspected malaria cases that received a laboratory diagnosis.

To evaluate the impact of scale-up of malaria control interventions, malaria-related morbidity should be measured from both population-based sources, such as household surveys, and routine data from health facilities.

VIII.1 Population-based Indicators

The Roll Back Malaria Monitoring and Evaluation Reference Group (RBM MERG) recommends that, in areas with high-intensity malaria transmission, anemia (defined as hemoglobin value less than 8g/dL) and malaria parasitemia prevalence should be measured in children ages 6–59 months as a sensitive and timely indicator of the morbidity impact of malaria control (see Table VIII.1).[1] Estimates of anemia and parasite prevalence are available from statistically sound and nationally representative, two-stage, sampled household surveys, such as Demographic Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and Malaria Indicator Surveys (MIS). These surveys typically are conducted every three to five years and provide national, and sometimes subnational, estimates for anemia and malaria parasite prevalence, as well as other health-related indicators of interest.

Here are some approaches to using survey data for morbidity analyses:

- Consider survey timing when comparing anemia and parasite prevalence among surveys. DHS and MICS often occur during the dry season (outside of the peak malaria transmission season), but MIS typically occur during the peak transmission season.
- Refine analyses by stratifying between urban and rural residence and region (especially if malaria endemicity varies across the country) and by wealth quintile as an indication of household vulnerability and equity in access, use of health services, and health outcomes.

- Tabulate anemia measurements as the prevalence of anemia with alternative hemoglobin thresholds, as well as the mean hemoglobin level with its standard deviation. This allows users to derive anemia prevalence for yet more alternative threshold levels by applying a normal approximation.[1,2]
- Malaria parasitemia often is determined microscopically or, as an alternative, with malaria rapid diagnostic tests (RDTs). When comparing results among surveys in settings of varying malaria endemicity, consider the varying test-specific sensitivity, specificity, and positive and negative predictive values; adjust interpretations accordingly.
- For both anemia and parasitemia, pay attention to the age cut-offs in measurements and comparisons. Malaria-related anemia and parasitemia peak at ages 1–2 years in high-endemic (Sub Saharan African) settings, but at higher ages in lower endemic settings. In particular, lower age cut-offs (such as 0, 1, or 6 months) markedly influence the prevalence.

Plasmodium parasite prevalence, an indicator of malaria morbidity, is the most commonly used metric to classify malaria endemicity levels. Several endemicity classifications are used to classify transmission intensity:

1. Method used by Malaria Atlas Project and Information for Malaria (INFORM)
   a. Define transmission risk, based on *P. falciparum* parasite rate in children ages 2 years up to 10 years (*Pf*PR 2–10) as follows:
      i. Low-risk PfPR 2–10, less than or equal to 5%
      ii. Intermediate risk PfPR 2–10, between 5% and 40%
      iii. High-risk PfPR 2–10, 40% or higher
   b. Then link global positioning system (GPS) location of household cluster with Malaria Atlas Project maps to identify the malaria transmission zone.
2. Methods that use Mapping Malaria Risk in Africa (MARA) project data
   a. Define transmission level as follows, based on 0–100% units of *Pf*PR 2–10:
      i. Low transmission, 0–20%
      ii. Medium transmission, 20–65%
      iii. High transmission, 65–100%
   b. Then link first-level administrative areas with MARA project data.
3. Methods used in the Tanzania impact evaluation report:
   a. Define three risk terciles that use malaria parasitemia in children ages 6–59 months:
      i. Lower tercile, less than 10%
      ii. Intermediate tercile, 10% to 20%
      iii. Higher tercile, over 20%
   b. Then link regions (about seven in each tercile) with MARA project maps.

TABLE VIII.1: MALARIA-RELATED POPULATION-BASED MORBIDITY INDICATORS

Indicator	Numerator	Denominator	Stratification	Frequency/Timing	Strengths/Limitations	Source
Proportion of children ages 6–59 months with malaria infection	Number of children with malaria parasitemia (microscopy or RDT)	Number of children tested for malaria parasitemia	Rural or urban residence, wealth quintiles, region	DHS: Every five years, dry season, low transmission season MIS: Every two to three years, high transmission season	This indicator directly measures parasite prevalence among children ages 6–59 months at the national level. However, parasite prevalence can fluctuate dramatically over a year with the seasonality of malaria, so survey timing may affect indicator values.	Survey (DHS, MIS)
Proportion of children ages 6–59 months with a hemoglobin measurement of less than 8 g/dL	Number of children with anemia	Number of children tested for anemia	Rural or urban residence, wealth quintiles, region		Provides a proxy measure of the prevalence of malaria-related anemia among children ages 6–59 months at the national level. However, this indicator is subject to seasonal variation in malaria-related anemia, which makes survey outcomes sensitive to the season of measurement.	Survey (DHS, MIS)

VIII.2 Health Facility-based Indicators

Malaria morbidity indicators based on health facilities include malaria outpatient visits or cases, hospital (inpatient) admissions, and, in young children in high-endemic settings, outpatient visits and hospitalizations for severe anemia (Table VIII.2). These indicators usually are obtained from surveillance and routine information systems, typically health management information systems (HMIS). These indicators should be stratified if possible by whether they are parasitologically confirmed (100% laboratory testing, 100% presumptive clinical diagnosis, or a mix). Typically, African countries have 100% laboratory confirmation among hospitalized cases, but not for outpatient visits.

To interpret outpatient data from settings with less than 100% or varying laboratory testing, evaluators must examine morbidity indicators alongside the malaria test positivity rate and the proportion of suspected malaria cases that have laboratory diagnosis. If coverage of service utilization and recording completeness is stable over the evaluation period, time trends in facility-based indicators still may provide an unbiased indication of the time trend in the total burden in the community. However, a change in case definition (that is, from clinical cases to laboratory-confirmed cases) or shortage in supply of RDTs or microscopy supplies could affect this time trend.

Changes in health care utilization and recording completeness can confound trends in the true malaria burden.[3] To account for this, it is important to assess the proportion of cases that malaria and, in stable endemic Sub Saharan Africa, anemia (when data are available) make up out of all-cause outpatient visits, hospital admissions, and hospital deaths.[4] The effect of malaria interventions, at least in stable endemic settings, may differ in children under 5 years of age and older children. Typically, the impact of malaria control on anemia is greater in children under 5 years of age than in children 5 years and older, among whom malaria is a less important contributor to anemia.

TABLE VIII.2: MALARIA-RELATED MORBIDITY INDICATORS FROM HEATH FACILITIES

Indicator	Numerator	Denominator	Stratification	Frequency and Timing	Strengths and Limitations
Outpatient malaria cases, annual rate	Number of confirmed malaria cases (by microscopy or RDT)	Catchment population of health facilities concerned, or in case of national HIS data, the national population at risk of malaria	All ages, children under age 5, sex, species, parasitologically confirmed, compared with all suspected fever cases. When available, prefer monthly data evaluated over several years	Information is collected routinely—daily, then compiled weekly, monthly, quarterly, and annually.	Provides information on outpatient level of malaria infection. However, change in use of health services and availability of diagnosis tests affects this indicator.
Outpatient malaria cases, proportion of all-cause outpatient cases	Number of confirmed malaria cases (by microscopy or RDT)	Total number of outpatient cases from any cause		Information is collected routinely—daily, then compiled weekly, monthly, quarterly, and annually.	Provides data on the contribution of malaria to outpatients compared with other illnesses. However, this indicator is sensitive to quality and availability of diagnosis tests.
Inpatient malaria cases and hospitalizations, annual rate	Hospitalizations with primary diagnosis of malaria at discharge	Persons at risk of malaria	All ages, children under age 5, sex	Information is collected routinely—daily, then compiled weekly, monthly, quarterly, and annually.	Provides information on the level of hospitalization due to malaria.
Inpatient malaria cases and hospitalizations, proportion of all-cause hospitalizations	Hospitalizations with primary diagnosis of malaria at discharge	Total number of hospitalizations from any case		Information is collected routinely—daily, then compiled weekly, monthly, quarterly, and annually.	Provides information on level of severe malaria compared with other causes of hospitalization. However, hospitalization capacity and use of health services affects this indicator.

Indicator	Numerator	Denominator	Stratification	Frequency and Timing	Strengths and Limitations
Malaria-test positivity	Number of outpatient laboratory-confirmed malaria cases	Total number of outpatient suspected malaria cases tested	All ages, children under age 5, sex, species, passive compared with active case detection, microscopy compared with RDT	Information is collected routinely—daily, then compiled weekly, monthly, quarterly, and annually.	Provides information of the level of infection among patients. However, this indicator is affected by variation in use of health services.
Proportion of suspected malaria cases with laboratory diagnosis	Number of suspected malaria cases that receive microscopy or RDT laboratory examination for malaria	Number of suspected malaria cases	All ages, children less than 5, sex	Information is collected routinely—daily, then compiled weekly, monthly, quarterly, and annually.	Provides information on the health capacity to test for malaria. However, this indicator is affected by availability of diagnosis commodities.

[1] Korenromp EL, Armstrong-Schellenberg JR, Williams BG, Nahlen BL, and Snow RW. Impact of malaria control on childhood anaemia in Africa—a quantitative review. Trop Med Int Health. 2004. 9(10):1050–1065.

[2] World Health Organization, Roll Back Malaria Department. Minutes of the First Meeting of the Roll Back Malaria MERG Task Force on Malaria-related Anemia. Geneva: World Health Institution. 2003. Available at: http://www.rollbackmalaria.org/partnership/wg/wg_monitoring/docs/MERG_Anaemia_tfm1_minutes.doc.

[3] Rowe AK, Kachur SP, Yoon SS, Lynch M, Slutsker L, et al. Caution is required when using health facility-based data to evaluate the health impact of malaria control efforts in Africa. Journal of Malaria. 2009. 8: 209.

[4] Aregawi M, Ali AS, Al-mafazy A, Molteni F, Warsame M, et al. Reductions in malaria and anemia case and death burden to hospitals following scale-up of malaria control in Zanzibar, 1999–2008. Malaria J. 2011. 10:46.

IX. Measuring Mortality

Key Points

- Malaria-specific mortality is challenging to measure at the population level because most deaths in malaria-endemic countries occur outside the health care system and without proper diagnostic tests.
- Routine health information systems provide useful data on all-cause and malaria-specific mortality at the subnational level. These data should be analyzed with a clear understanding of limitations.
- Civil registration and vital statistics systems provide information on all-cause mortality and, in some cases, malaria-specific mortality. However, for most malaria endemic countries coverage is too low to provide unbiased estimates at the national level.
- Health and demographic surveillance systems provide longitudinal data on malaria-specific mortality, but these data are restricted to specific geographic areas within countries.
- Verbal autopsy tools provide data on malaria-specific mortality at the population level and can be nested to other data sources. However, these tools have low specificity and sensitivity in detecting malaria mortality.
- Given the limited availability of malaria-specific mortality data, the MERG recommends using all-cause child mortality as an impact indicator. This can be measured reliably at the national level through household surveys.

Measuring malaria-specific mortality and all-cause mortality at the national level for the purpose of assessing the impact of malaria control interventions requires high-quality data. Several potential data sources, each with its own set of limitations, can provide the required data. These sources include routine health management information systems (RHMIS), civil registration and vital statistics (CRVS), health and demographic surveillance systems (HDSS), verbal autopsy (VA), household surveys, and national censuses.

IX.1 Malaria-specific Mortality

IX.1.1 Routine Health Management and Information Systems and Sentinel Surveillance

One measure of malaria mortality that is available in nearly all malaria-endemic countries is deaths attributed to malaria among cases that are recorded in RHMIS, either in general health reporting or as part of separate reporting for malaria. Because malaria deaths are likely to be derived from deaths among inpatient malaria cases, this information may be available even if malaria-attributable deaths have not been recorded as a separate indicator in monthly facility, district, or regional reports. It is difficult to measure malaria-specific mortality with reasonable precision because of several factors

that range from malaria's large indirect contribution to mortality to reliability of malaria diagnosis and the variable quality of health reporting. However, information on malaria deaths reported through RHMIS can complement other sources of malaria mortality by addressing limitations in the other systems.

Evaluating other measures of malaria morbidity in RHMIS can enhance the tracking of trends in malaria deaths. Evaluators should consider several types of variables available in HMIS along with malaria deaths to assess trends in malaria mortality to see if they are consistent with other trends in malaria morbidity. Malaria cases, both clinically defined and confirmed by diagnostic testing, often will follow trends similar to malaria deaths. Inpatient malaria cases are likely to be more severe and strongly correlate with malaria mortality. Tracking the proportion of all deaths due to malaria may distinguish between true changes in malaria epidemiology and changes in facility utilization.

Advantages of Using RHMIS Data to Assess Malaria Mortality

Assessing malaria mortality in RHMIS offers certain advantages over other sources of malaria-related mortality, such as all-cause child mortality (ACCM) in household surveys; verbal–autopsy-defined, malaria-specific deaths; and malaria deaths from vital registration systems. Routine system health data often are available at the district level, and they are widely available in nearly all countries and recorded monthly in most places. Because malaria information is part of an overall routine health information system, the resources required to access the system are modest and sustainable over time. Malaria-specific mortality information reported from health facilities through RHMIS is often among cases of laboratory-confirmed malaria. Although RHMIS can miss deaths that result indirectly from malaria, RHMIS can likely offer greater specificity on malaria deaths than other sources because these deaths occur in facilities.[1] Most of these systems include information on age and sex, and therefore, malaria deaths reported in RHMIS can be stratified by these factors to obtain more detailed trends. RHMIS-reported malaria deaths reflect the burden malaria places on a country's health care system, an important component in the overall assessment of a country's malaria burden. RHMIS data, under the control of the ministry of health, may provide timely information for district-level program management.

Disadvantages of Using RHMIS Data to Assess Malaria Mortality

The disadvantages of using RHMIS data to assess malaria mortality are commonly cited about routine health information systems. The quality of malaria diagnoses of inpatient malaria cases can be poor and may vary over time. Incomplete recording of malaria inpatient cases and deaths at health facilities and inconsistent reporting to RHMIS also are problematic. Trends in health care utilization affect malaria deaths at health facilities. Low utilization, coupled with the fact that some malaria-related deaths occur at private facilities or in homes, means only a fraction of malaria deaths in a community are likely to occur at public health facilities that report deaths through the RHMIS.[2]

Deaths that do occur in public facilities may not be representative of all malaria deaths, which may pose challenges for disaggregating by demographic or other characteristics.

Considerations for Analysis

An analysis of trends in malaria-specific mortality based on RHMIS data that optimizes reported data from health facilities and considers data limitations can yield valuable results for malaria programs. In addition to influencing factors, evaluators should consider that some changes in malaria control programming can affect reported data. Continued scale-up of diagnostic testing will improve the quality of data, but large and rapid changes in the proportion of confirmed cases may make it difficult to assess trends in confirmed malaria deaths alone. In this case, an analysis could examine the proportion of all heath facility malaria deaths to yield a more reliable measure of trends in malaria mortality. As malaria control and treatment improve, inpatient malaria fatality rates also may change, which evaluators should consider when reviewing trends in overall malaria mortality. On the other hand, as malaria programs receive more support, health facilities may increase the rate they report to HMIS, which could artificially increase the reported number of malaria deaths.

Box IX.1: Civil Registration and Vital Statistics

Civil registration and vital statistics systems (CRVS) have the potential to provide direct measurement of population-level mortality for malaria, if malaria diagnosis is reliable and medical autopsies are performed. Trends in all-cause, age-specific mortality recorded in CRVS systems suggest that it is possible to track the impact of scaling up antiretroviral treatment on HIV-attributable deaths.[3,4,5] CRVS-recorded deaths are starting to produce informative time trends on tuberculosis in an increasing number of low- and especially middle-income countries.[6,7] Unfortunately, in most low- and middle-income countries where malaria and other major infectious diseases are concentrated, CRVS systems function poorly and produce incomplete data.

Only a few countries have been able to improve their CRVS systems significantly in the past 50 years.[8] For most malaria-endemic countries, notably in Sub Saharan Africa, the first challenge is to build and roll out a national-scale functioning CRVS system that captures all deaths.[9] None of the malaria-endemic countries in recent national-level estimations of the causes and time trends in child mortality used CRVS data.[10,11] Studies in a few malaria-endemic countries have tried to validate the sensitivity and specificity of CRVS for malaria-attributable deaths, comparing CRVS-based, malaria-death attribution with the results of verbal autopsy (Box VIII.2) in health and demographic surveillance sites or in sample civil registration and vital statistics systems[12,13,14,15] or routine hospital records.[16] These studies suggested poor levels of agreement between CRVS and the CRVS- or hospital-based recordings, which results in major underreporting of malaria-attributable deaths in children in rural Kenya[15] or overreporting on children and adults in Tanzania.[16] The ongoing rollout of malaria parasitological diagnostic tools (microscopic or rapid diagnostic tests) as part of routine malaria case management is expected to improve the reliability of malaria-death recording in hospitals.

Political commitment has increased in recent years, as demonstrated by the Africa Programme on Accelerated Improvement of Civil Registration and Vital Statistics and endorsed by African Ministers in Ethiopia (2010) and South Africa (2012) and the Global Summit on Civil Registration and Vital Statistics (2013). The Millennium Development Goals Africa Steering Group estimated a cost of US $80 million, or US $0.10 per person, to achieve CRVS in Africa. This is generally considered a reasonable price for a global public health good that improves the progress of monitoring and yields an actual impact from health investments.[17,18] Some countries, such as Ethiopia and Rwanda, have started to use recent increased funding for health to employ and empower community health workers throughout the country to record deaths and births in communities. This type of initiative has the potential to improve monitoring all-cause, under-5 mortality and may enhance the reporting of deaths for specific selected causes.[17]

CRVS systems are important to develop over the long term for general health and development. These systems also can serve as a tool for strengthening national health systems, setting priorities, and evaluating disease programs.[19] As malaria control expands and malaria declines as a leading cause of death in young children, CRVS systems could become a more useful data source for tracking malaria-attributable deaths. National governments and global health initiatives, including Roll Back Malaria, should call for improving CRVS systems, and urge donors and international funding agencies to support this.[19]

IX.1.2 Health and Demographic Surveillance Systems

As a response to the lack of reliable vital registration in low- and middle-income countries, a number of developing countries (mostly in Africa and Asia) have set up Health and Demographic Surveillance Systems (HDSSs) to record longitudinal demographic data in geographically defined areas. The International Network for the Demographic Evaluation of Populations and Their Health in Developing Countries (INDEPTH-network) coordinates 44 HDSSs, 32 of which are in 19 African countries (www.indepth-network). Most HDSSs begin with a baseline census of the population under surveillance. Subsequently, this population is monitored at regular intervals, ranging from one to four times a year, to record vital events (births, deaths, and migrations). The population under surveillance varies from site to site. Trained interviewers or local registrars conduct the surveillance through household visits. Deaths recorded by interviewers are followed up with verbal autopsy interviews to gain cause-specific mortality data. Increasingly, HDSSs have received attention for their ability to provide mortality data in countries that lack a CRVS system.[20,21]

For much of the developing world, HDSSs have the highest potential as sources of data on cause-specific mortality. Data collected from HDSSs have contributed to an understanding of the mortality burden in defined populations in developing countries.[22,23,24,25] Officials have successfully used mortality data from HDSSs for health planning in resource-constrained settings. For example, in Tanzania, officials used mortality data to plan district and national health programs; the burden of disease profiles helped set priorities, which resulted in a measurable decline in child mortality.[26] Data from HDSSs can help estimate regional and global disease burdens,[27,28,29] and HDSSs play a crucial role in producing data for health planning and management in countries that lack routine CRVS systems.[30]

Collecting data in HDSSs also has some limitations. One major criticism is that data are collected from relatively small, defined populations; therefore, the data may not be applicable to other populations. To overcome this limitation, some countries set up HDSSs in locations that represent major climatic or geographic settings.[31] The high operational costs of running these sites are also a hindrance, although some have argued that the per-capita cost of running an HDSS is less than operating a routine HMIS.[32] By their nature as sentinel sites, HDSS-covered populations often benefit from above-average health care. This could cause mortality rates in HDSSs to differ from those of the broader African population, although the proportional distribution of causes of deaths is probably less biased.[27]

Box IX.2: Verbal Autopsy

Verbal autopsy (VA) is a method of collecting data on the probable causes of death. Demographic Health Surveys, health and demographic surveillance systems, and civil registration and vital statistics systems all use the VA method. VA involves interviewing primary caregivers of recently deceased people to gather information on the circumstances of death by using standard data-collection instruments determined by the World Health Organization and its partners.[33] Interpreting information on the circumstances of death helps derive probable causes of death, using the International Classification of Diseases Version 10 (ICD-10) list or an abridged version.[34,35,36,37,38]

VA provides an alternative method to record causes of death in settings where CRVS is not functioning. Although VA works better at identifying some causes of death than others, the consensus is that it remains the best alternative to use in developing countries without a national CRVS to understand cause-of-death patterns.9[33,39] SCRVS, which continuously register births and deaths in a representative sample of population clusters throughout a country, can include VA.[40,41,42] In this situation, causes are classified based on verbal autopsies for deaths that occur at home, which is typically a large proportion of total deaths.

Based on information collected during verbal autopsies, officials use several methods to ascertain and summarize causes of death. These include a physician's review, expert algorithms, data-derived algorithms, and computer-based modeling (probabilistic and symptom-pattern methods).[35,43,44,45,46]

Concerns about the overall validity of verbal autopsy data mostly pertain to a lack of standardization of collection methods[47] and the variable sensitivity and specificity across diseases. VA appears to work well to classify deaths resulting from certain diseases of public health significance (such as measles, whooping cough, and cholera), because of the distinct symptoms and signs of these conditions. Verbal autopsies also can help classify deaths that result from injuries and violence. Verbal autopsies, however, are not as sensitive for conditions with less specific symptoms and signs, such as malaria.[48] Despite the imperfect sensitivity, VA has an important role to play in providing information on malaria-specific mortality. A number of studies that used verbal autopsy data quantified the burden and trends in malaria mortality.[49,50,51,52] For example, a study of a town in northwestern Burkina Faso between 1999 and 2004 demonstrated that all-cause mortality rates declined, while malaria-attributed mortality remained constant.[53] Another study in Ethiopia confirmed the seasonality of malaria mortality and treatment patterns and care-seeking behavior of malaria victims before death.[50]

VA validation studies conducted in malaria-endemic settings have estimated low sensitivity (24 to 75%) and moderate to high specificity (77 to 100%) for malaria; in other words, verbal autopsies often miss true malaria deaths, but conversely misclassify non-malaria deaths as malaria.[37,54] The reason for the low sensitivity may be clinical presentation, especially in children under age 5 years, who are affected the most. Because of their underdeveloped immune systems, these children present nonspecific symptoms, such as fever, vomiting, and diarrhea. If the sensitivity and specificity have been estimated in local validation studies, analysts can adjust for malaria misclassification in verbal autopsies.

Malaria-specific mortality data derived from verbal autopsies may have validity issues, especially in relation to their low sensitivity. However, considering the poor state of CRVS systems in most of the countries that have the highest burden of malaria, verbal autopsies provide a reasonable short-to-medium-term source of malaria-specific mortality data. It is difficult to point to any of the platforms that use verbal autopsies—such as HDSS, DHS, census, or SCRVS—as the most ideal to generate malaria-specific mortality because of their limitations. Combining data from different systems can help derive better estimates. For example, evaluators could combine hospital-based data on malaria mortality with data gained through verbal autopsies from an HDSS, using the symptom pattern method to generate population-level malaria-specific mortality fractions; the precise approach or combination of approaches used depends on the available context-specific resources. As much as possible, evaluators should use standardized processes for verbal autopsies, questionnaires, and analysis tools to ensure high-quality data.

Timely, reliable mortality data by age, sex, and cause—both nationally and subnationally—are critical for the design, implementation, and monitoring and evaluation of health programs. Countries should capitalize on VA strengths to generate reasonable estimates of malaria-specific mortality data rather than let the limitations of verbal autopsies hinder them.

IX.1.3 Demographic and Health Surveys and National Censuses

Demographic and Health Surveys

DHS are conducted every five years on nationally representative samples of between 3,000 and 30,000 households (www.measuredhs.com). DHS results contribute the major source of health data for many low-income countries.[31] Attempts have been made recently in some countries—Ghana, Rwanda, and Afghanistan—to incorporate verbal autopsies into DHS. Households with a recorded death during the survey are revisited to complete the verbal autopsy questionnaire. In Ghana during the 2008 DHS, verbal autopsies were performed for deaths of children under age 5 years,[55] but were limited to those that occurred in the past three years from the date of the interview.

Several limitations affect DHS data, including data gained in verbal autopsies. The surveys are prone to recall bias because they only are conducted once every three to five years. Primary caregivers can find it difficult to provide accurate information on the circumstances of a death that occurred more than a few months before the DHS interview. Verbal autopsies are, therefore, limited to deaths that occurred in the last two years. Furthermore, for logistical reasons, DHS surveys are usually conducted during weather-friendly seasons, which might affect the estimation of causes of death that resulted from seasonal conditions, such as malaria. DHS samples are not powered to generate estimates at the lowest administrative level, such as districts or counties, which limits the usefulness of DHS verbal autopsy data for local health planners.[32]

National Censuses

National censuses, which provide vital information on the total population of countries, offer a possibility to incorporate verbal autopsies. Mozambique has experimented with this method in its Post-Census Mortality Survey.[56] Essentially, they selected a sample of census enumeration areas—representative at the national, urban-rural, and provincial levels—which formed a so-called "death frame." All households that reported a death in the sample enumeration areas received a follow-up with a verbal autopsy form. A team of trained doctors reviewed the verbal autopsies and assigned a cause-of-death code following ICD-10 procedures. The main limitation of including verbal autopsies in censuses is that the resulting estimates are of average mortality over a long period in the past; therefore, they may underestimate overall mortality. As with DHS, recall bias also may be a limitation.

IX.1.4 Ways Forward—Examine, Improve Data Quality and Completeness

Although HMIS limitations can be accounted for during analysis, perhaps the best way to improve measurement of malaria mortality is to increase the quality of the reporting system itself. Not all deaths among inpatients that demonstrate parasites in a diagnostic test are a result of malaria. Some deaths may be indirectly attributable to malaria, although they show no evidence of current

infection. Still, the number of deaths among patients with laboratory-confirmed malaria offers a high level of specificity in measuring malaria mortality that is not available from other data sources. As such, scale-up of diagnostic testing for malaria among inpatients, through both microscopy and RDTs, is likely to improve mortality measurement among facility users.

Analyses can include all facilities if they have consistently high reporting and data completeness. If there are quality concerns from certain facilities, evaluators may remove them from analyses to prevent biased results. If there are quality concerns from most facilities or if completeness and consistency of data cannot be assessed adequately, evaluators may find it necessary to collect data directly from selected health facilities through a rapid impact assessment.[57] With this approach, a data collection team systematically samples facilities, visits them, and reviews registers and other records for relevant data on malaria and other conditions to consider completeness and data quality.

The analysis should consider several variables (such as health care utilization) that can affect trends in malaria-related deaths reported in HMIS. A comparison of trends in malaria inpatient cases to trends in nonmalaria inpatient cases, stratified by age, may help assess if health care utilization for malaria has changed. As with other time series data, trends in facility-based malaria deaths should be adjusted for changes between the pre- and post-intervention periods. A comparison of the percentage of deaths that occur in facilities to all deaths in the country can give a sense of the magnitude of bias in using facility data to estimate all malaria-related deaths. If the ratios of causes of death between facilities and a community are consistent, researchers might derive a correction factor to apply to trends in facility malaria deaths from HMIS.

IX.2 All-Cause Child Mortality from Household Surveys

In malaria-endemic countries, malaria is a key contributor to all-cause child mortality (ACCM), accounting for an estimated 15% of all under-5 deaths in Sub Saharan Africa.[58] It is difficult to measure malaria-attributable mortality with any precision, in part because of malaria's large indirect contribution to increasing the risk of death from other common childhood illnesses, such as pneumonia, diarrhea, anemia, and measles;[59] therefore, RBM MERG recommends using ACCM as the standard measure of impact in malaria-endemic countries.

The United Nations and other major donor organizations also use ACCM to track global commitments. It is a standard indicator of disease burden and one of the primary indicators for Millennium Development Goals (MDGs), especially (4) Reduce Child Mortality, with the target being a two-thirds reduction in all-cause, under-5 mortality from 1990 to 2015. The Global Fund to Fight AIDS, Tuberculosis and Malaria uses ACCM as one of its corporate key performance indicators for country-level health impact; it recommends this measure as an essential impact indicator for malaria grants to report on at periodic review for grant renewal.[60] National governments use ACCM to monitor the performance of the overall health sector.

Researchers commonly measure ACCM through standardized household surveys, such as DHS and MICS. They derive estimates of the number of deaths for children under age 5 years from a series of questions on the full birth histories of women interviewed. DHS and MICS surveys typically are undertaken every three to five years, and researchers interpret the under-5 mortality estimate from a survey to pertain to the time about 2.5 years before the survey because birth histories cover the past five years. The peak age in under-5 mortality is 1 to 2 years. DHS and MICS reports—including their estimates of infant, child, and overall under-5 mortality by urban and rural stratification—are published within an average of six to nine months following the survey. Most data are available in the public domain for use by program managers, researchers, and others.

Most developing countries primarily obtain under-5 mortality estimates from household surveys; researchers should consider the attached confidence intervals when comparing values over time or across countries. Non-sampling errors that may affect equally recent levels and trends often influence these survey estimates. The Inter-agency Group for Child Mortality Estimation (IGME)—which comprises the United Nations' Children Fund (UNICEF), WHO, the World Bank, and the United Nations Population Division, as well as independent technical experts—formed in 2004. IGME's aim is to improve methods for child mortality estimation and produce consistent country-level estimates and trends of child mortality worldwide. IGME estimates are used to report on progress toward MDGs. IGME updates its under-5 mortality estimates after reviewing all newly available country-level data and assessing data quality. IGME methodology minimizes errors in data quality and maximizes the consistency of trends observed during the past 30 years. IGME estimates are produced once a year and released every September at the national level only. IGME plans to produce estimates that are disaggregated by sex and urban-rural residence. Country-specific estimates and the data used to derive them are available from IGME's child mortality database, CME Info, at www.childmortality.org. IGME's full methodological details are available at www.childinfo.org.

Although large national surveys are generally the best sources of data to calculate ACCM in settings with limited data on mortality, some countries collect information from other sources. The other most reliable data source on ACCM is a national census—the gold standard for mortality data—when it is available. Censuses typically are conducted every 10 years, although some countries in Sub Saharan Africa have not conducted a census for much longer than that. Evaluators should use new census data, when available, to estimate ACCM instead of survey data. The IGME group adds new census data to its calculations when it becomes available.

A second source of data is national vital registration systems. As discussed earlier, vital registration systems in many countries in Sub Saharan Africa suffer from large variations in the completeness and coverage of information. In some cases, evaluators can use vital registration systems for ACCM estimations following statistical adjustments to correct for incompleteness. Analysts also can use them at the subnational level in areas (often urban settings) where coverage is better; however, the

variability of these data sources makes them less desirable for large cross-national comparisons or for looking at trends over time.

Many countries in Africa have HDSSs that collect data on both ACCM and cause-specific mortality. Data from these sites are useful to calculate the case-specific ACCM, and the data tend to be of high quality. HDSSs usually investigate every death in the site, often using verbal autopsy methods so that cause of death information is generalizable to the larger community. In addition to cause-of-death data, HDSSs conduct censuses in their communities, which yield ACCM data for that particular site. HDSSs are, by definition, research sites that are subject to a range of biases because they are under frequent observation. These sites also frequently benefit from pilots of health-related interventions, meaning morbidity and mortality data are not necessarily nationally representative and must be interpreted with caution.

MERG recommendations for using ACCM as an impact indicator specify that it is only for highly endemic settings. To be consistent with other global initiatives, IGME estimates under-5 mortality for impact evaluation. However, because of the limited possibilities for disaggregation, other sources of data (such as DHS and MICS, if available) should be used for country-level regression analyses.

---

[1] Oti SO, and Kyobutungi C. Verbal autopsy interpretation: a comparative analysis of the InterVA model versus physician review in determining causes of death in the Nairobi DSS. *Population Health Metrics*. 2010. 8:21.

[2] World Malaria Report. Geneva, World Health Organization, 2013.

[3] Herbst A, Mafojane T, and Newell M. Verbal autopsy-based cause-specific mortality trends in rural KwaZulu-Natal, South Africa, 2000-2009. *Population Health Metrics*. 2011. 9:47.

[4] Stoneburner R, Korenromp E, Lazenby M, Tassie J-M, Letebele J, et al. Using Health Surveillance Systems Data to Assess the Impact of AIDS and Antiretroviral Treatment on Adult Morbidity and Mortality in Botswana. *PLoS ONE*. 2014. 9(7): e100431.

[5] Rhucharoenpornpanich O, et al. Use of vital statistics in estimating HIV prevention and treatment effects in Thailand. 19th International AIDS Conference. Abstract no. LBPE58.

[6] Korenromp EL, Bierrenbach AL, Williams BG, and Dye C. The measurement and estimation of tuberculosis mortality. *Int J Tuberc Lung Dis*. 2009. 13:283–303.

[7] Glaziou P, Floyd K, Korenromp E, et al. Lives saved by tuberculosis control and prospects for achieving the 2015 global target for reductions in tuberculosis mortality. *Bull World Health Organ*. 2011. 89(8):573–582.

[8] Mahapatra P, Shibuya K, Lopez AD, et al. Civil registration systems and vital statistics: Successes and missed opportunities. *Lance*t. 2007. 370(9599):1653–1663.

[9] Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: An assessment of the global status of cause of death data. *Bull World Health Organ*. 2005. 83(3):171–177.

[10] Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: A systematic analysis. *Lancet*. 2010. 375(9730):1969–1987.

[11] Rajaratnam JK, Marcus JR, Flaxman AD, et al. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970–2010: A systematic analysis of progress towards Millennium Development Goal 4. *Lancet*. 2010. 375(9730): 1988–2008.

[12] Jha P, Gajalakshmi V, Gupta PC, et al. Prospective study of one million deaths in India: Rationale, design, and validation results. *PLoS Med*. 2006. 3:e18.

[13] Mung'ala VO, and Snow RW. Death registration on the Kenyan Coast. *East Afr Med J*. 1994. 71(11):747–750.

[14] Edward A, Ernst P, Taylor C, Becker S, Mazive E, and Perry H. Examining the evidence of under-five mortality reduction in a community-based programme in Gaza, Mozambique. *Trans R Soc Trop Med Hyg*. 2007. 101(8):814–822.

[15] Arudo J, Gimnig JE, ter Kuile FO, et al. Comparison of government statistics and demographic surveillance to monitor under-five child mortality in rural western Kenya. *Am J Trop Med Hyg*. 2003. 68(4 Suppl):30–37.

[16] Reyburn H, Mbatia R, Drakeley C, et al. Overdiagnosis of malaria in patients with severe febrile illness in Tanzania: A prospective study. *BMJ*. 2004. 329(7476):1212.

[17] Zachariah R, Bemelmans M, Akesson A, Gomani P, Phiri K, Isake B, et al. Reduced tuberculosis case notification associated with scaling up antiretroviral treatment in rural Malawi. *Int J Tuberc Lung Dis*. 2011. 15(7):933–937.

[18] Achieving the Millennium Development Goals In Africa: Recommendations of the MDG Africa Steering Group, June 2008. New York: United Nations Department of Public Information, 2008.

[19] Chan M, Kazatchkine M, Lob-Levyt J, et al. Meeting the demand for results and accountability: A call for action on health data from eight global health agencies. *PLoS Med*. 2010. 7:e1000223.

[20] Baiden F, Hodgson A, Binka FN. Demographic surveillance sites and emerging challenges in international health. *Bull World Health Organ*. 2006. 84(3):163.

[21] Byass P, Berhane Y, Emmelin A, Kebede D, Andersson T, Högberg U, et al. The role of demographic surveillance systems (DSS) in assessing the health of communities: An example from rural Ethiopia. *Public Health*. 2002. 116(3):145–150.

[22] Kyobutungi C, Ziraba AK, Ezeh A, and Yé Y. The burden of disease profile of residents of Nairobi's slums: Results from a Demographic Surveillance System. *Popul Health Metr*. 2008. 6:1.

[23] van Eijk AM, Adazu K, Ofware P, Vulule J, Hamel M, and Slutsker L. Causes of deaths using verbal autopsy among adolescents and adults in rural western Kenya. *Trop Med Int Health*. 2008. 13(10):1314–1324.

[24] Herbst AJ, Cooke GS, Barnighausen T, KanyKany A., Tanser F, and Newell ML. Adult mortality and antiretroviral treatment rollout in rural KwaZulu-Natal, South Africa. *Bull World Health Organ*. 2009. 87(10):754–762.

[25] Jahn A, Floyd S, Crampin AC, Mwaungulu F, Mvula H, Munthali F, et al. Population-level effect of HIV on adult mortality and early evidence of reversal after introduction of antiretroviral therapy in Malawi. *Lancet*. 2008. 371(9624):1603–1611.

[26] de Savigny D, Kasale H, Mbuya C, *Reid G. In focus: Fixing Health Systems*. 2005. Ottawa: International Development Research Centre.

[27] Korenromp EL, Williams BG, Gouws E, Dye C, and Snow RW. Measurement of trends in childhood malaria mortality in Africa: An assessment of progress toward targets based on verbal autopsy. *Lancet Infect Dis*. 2003. 3(b):349–358.

[28] Morris SS, Black RE, and Tomaskovic L. Predicting the distribution of under-5 deaths by cause in countries without adequate vital registration systems. *Int J Epidemiol*. 2003. 32(6):1041–1051.

[29] INDEPTH. *INDEPTH Monograph Series: Population, Health and Survival at INDEPTH Sites*. 2002. Ottawa: IDRC.

[30] Ye Y, Wamukoya M, Ezeh A, et al. Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Sahara Africa? *BMC Public Health*. 2012. 12:741.

[31] Hill K, Lopez AD, Shibuya K, and Jha P. Interim measures for meeting needs for health sector data: Births, deaths, and causes of death. *Lancet*. 2007. 370(9600):1726–1735.

[32] de Savigny D, Mayombana C, Mwageni E, et al. Care-seeking patterns for fatal malaria in Tanzania. *Malaria Journal*. 2004. 3:27.

[33] Baiden F, Bawah A, Biai S, Binka F, Boerma T, Byass P, et al. Setting international standards for verbal autopsy. *Bull World Health Organ*. 2007. 85(8):570–571.

[34] Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: Current practices and challenges. *Bull World Health Organ.* 2006. 84(3):239-245.

[35] Quigley MA, Chandramohan D, and Rodrigues LC. Diagnostic accuracy of physician review, expert algorithms and data derived algorithms in adult verbal autopsies. *Int J Epidemiol.* 1999. 28(6):1081–1087.

[36] Kahn K, Tollman SM, Garenne M, and Gear JS. Validation and application of verbal autopsies in a rural area of South Africa. *Trop Med Int Health.* 2000. 5(11):824–31.

[37] Setel PW, Whiting DR, Hemed Y, Chandramohan D, Wolfson LJ, Alberti KG, and Lopez AD. Validity of verbal autopsy procedures for determining cause of death in Tanzania. *Trop Med Int Health.* 2006. 11(5):681–696.

[38] Gajalakshmi V, Peto R, Kanaka S, and Balasubramanian S. Verbal autopsy of 48,000 adult deaths attributable to medical causes in Chennai (formerly Madras), India. *BMC Public Health.* 2002. 2:7.

[39] Ye Y, Wamukoya M, Ezeh A, Emina JB, Sankoh O. Health and demographic surveillance systems: A step towards full civil registration and vital statistics system in sub-Sahara Africa? *BMC Public Health.* 2012. 12:741.

[40] Bassani DG, Kumar R, Awasthi S, et al. Causes of neonatal and child mortality in India: A nationally representative mortality survey. *Lancet.* 2010. 376(9755):1853–60.

[41] Dhingra N, Jha P, Sharma VP, et al. Adult and child malaria mortality in India: A nationally representative mortality survey. *Lancet.* 2010. 376(9754):1768–74.

[42] Rao, C et al. Evaluating national cause-of-death statistics: principles and application to the case of China. *Bull World Health Organ.* 2005. 83(8): 618–625.

[43] Soleman N, Chandramohan D, and Shibuya K. Verbal autopsy: Current practices and challenges. *Bull World Health Organ.* 2006. 84: 239–245.

[44] Boulle A, Chandramohan D, and Weller P. A case study of using artificial neural networks for classifying cause of death from verbal autopsy. *Int J Epidemiol.* 2001. 30(3):515–520.

[45] Byass P, Huong DL, and Minh HV. A probabilistic approach to interpreting verbal autopsies: Methodology and preliminary validation in Vietnam. *Scand J Public Health.* 2003. 31(Suppl.62):32–37.

[46] Murray CJL, Lopez AD, Feehan DM, Peter ST, and Yang G. Validation of the symptom pattern method for analyzing verbal autopsy data. *PLoS Medicine.* 2007. 4(11):e327.

[47] Joshi R, Kengne AP, and Neal B. Methodological trends in studies based on verbal autopsies before and after published guidelines. *Bull World Health Organ.* 2009. 87(9):678–682.

[48] Garenne M, and Fauveau V. Potential and limits of verbal autopsies. *Bull World Health Organ.* 2006. 84(3):164.

[49] Korenromp EL, Williams BG, Gouws E, Dye C, and Snow RW. Measurement of trends in childhood malaria mortality in Africa: An assessment of progress toward targets based on verbal autopsy. *Lancet Infect Dis.* 2003. 3(b):349–358.

[50] Deressa W, Fantahun M, and Ali A. Malaria-related mortality based on verbal autopsy in an area of low endemicity in a predominantly rural population in Ethiopia. *Malar J.* 2007. 6:128.

[51] Ndugwa RP, Ramroth H, Muller O, Jasseh M, Sie A, Kouyate B, et al. Comparison of all-cause and malaria-specific mortality from two West African countries with different malaria transmission patterns. *Malar J.* 2008. 7:15.

[52] O'Meara WP, Bejon P, Mwangi TW, Okiro EA, Peshu N, Snow RW, et al. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. *Lancet.* 2008. 372(9649):1555–1562.

[53] Ramroth H, Ndugwa RP, Muller O, Ye Y, Sie A, Kouyate B, et al. Decreasing childhood mortality and increasing proportion of malaria deaths in rural Burkina Faso. *Glob Health Action.* 2009. 2.

[54] Rowe AK. Should verbal autopsy results for malaria be adjusted to improve validity? *Int J Epidemiol.* 2005. 34(3):712-713.

[55] Ghana Statistical Service (GSS), GHSG, ICF Macro. *Ghana Demographic and Health Survey 2008*. 2009. Accra, Ghana: GSS, GHS, and ICF Macro.

[56] Instituto Nacional de Estatistica, 2007 unpublished data.

[57] Habicht J, Victora C, and Vaughan J. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol.* 1999. 28(1):10–18.

[58] Liu L, Johnson HL, Cousens S, et al. Global, regional, and national causes of child mortality: An updated systemic analysis for 2010 with time trends since 200. *Lancet.* 379(9832): 2151–2161.

[59] Rowe AK, Steketee RW, Arnold F, Wardlaw T, Basu S, Bakyaita N, et al. Viewpoint: Evaluating the impact of malaria control efforts on mortality in sub-Saharan Africa. *Trop Med Int Health*. 2007. 12(12):1524–1539.

[60] The Global Fund to Fight AIDS Tuberculosis and Malaria. *The Global Fund Results Report 2011: Making a Difference.* 2011. Geneva.

X. Measuring and Accounting for Nonmalaria Programs, Contextual Factors

Key Points

- Contextual factors can (1) confound the association between scale-up of the intervention and its potential health impact and (2) modify the effect of the intervention, and thus, affect the conclusion.
- Contextual factors differ with variations in data quantity and quality in and among countries. Each country evaluation should list relevant contextual factors and assess whether to change them and to what degree they can affect outcomes on the mortality of children under age 5 years.
- Although contextual factors vary from country to country, all countries must assess some core contextual factors.

Appropriate consideration of nonmalaria programs and other contextual factors is essential to ensure internal and external validity of evaluations of large-scale health programs that vary substantially in their impact.[1] Contextual factors can (1) confound the association between scale-up of the intervention and its potential health impact and (2) modify the effect of the intervention, and thus, affect the conclusion.

To evaluate thoroughly the relationship between the scale-up of malaria control interventions and decreases in mortality of children under age 5 years, it is essential to determine if contextual factors offer possible explanations for the observed mortality reductions. Table X.1 lists examples of contextual factors to consider during an evaluation of malaria control programs. It is likely that, in and among countries, influencing contextual factors differ with variations in data quantity and quality. At a minimum, each country evaluation should list relevant contextual factors and assess whether and to what degree changes in these factors could have affected child mortality.

TABLE X.1: 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, such as vitamin A, iron, zinc	WHO, UNICEF annual estimates of national immunization coverage UNICEF Vitamin A coverage database DHS, MICS, MIS	Observed reductions in child morbidity and mortality may actually be due to increased coverage of these programs rather than malaria control interventions.
Climatic and environmental factors	Total precipitation Number of days with rain Land cover and vegetation Air temperature Extreme weather events, such as floods	National meteorological agency Columbia University Earth Institute National Oceanographic and Atmospheric Administration	These factors affect mosquito breeding and malaria transmission and may cause observed changes in outcomes over time or geography, rather than the interventions themselves.
Health systems factors	Per capita expenditure on health Government expenditure on health as percentage of total government expenditure Availability of essential drugs Political situation and stability	WHO The World Bank	Health systems can affect comparisons across time or geography by influencing access to interventions. These factors modify the impact of malaria control interventions.
Socioeconomic factors	Household assets and income Parental education Conflict or emergency settings GDP per capita Gini per capita Population living below poverty line	DHS, MICS The World Bank	If different socioeconomic groups access malaria control interventions differently, these factors may serve as effect modifiers that influence outcomes.

Notes: 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 Survey; MIS=Malaria Indicator Survey; GDP=Gross domestic product.

X.1 Data Sources

The Demographic Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and Malaria Indicator Surveys (MIS) include detailed questions on nonmalaria programs and factors of interest, including socioeconomic status, health and nutritional indicators, coverage of health care services, and immunization coverage. Other sources of these indicators are the World Health Organization (WHO) and World Bank reports, United Nations Children's Fund (UNICEF), and country-specific reports. Additional sources may include other country-specific surveys and datasets identified in each country's data collection process, such as Health and Demographic Surveillance Systems (HDSSs) data through discussions with the Ministry of Health (MoH), the National Malaria Control Program (NMCP), in-country nongovernmental organizations, and other partners with knowledge of the country's health systems.

The analysis may include data from sources other than DHS, MIS, and MICS under the following circumstances: the survey methodology is available for review to determine the sampling scheme and representativeness, data are stratified or age information is available so that results for children under age 5 years can be extracted, survey sampling weights are available for secondary analysis of existing data, and less than 5% of data are missing for indicators of interest.

Datasets available online, containing information for multiple countries over many years, include these examples:

- DHS and MIS datasets, at http://www.measuredhs.com/
- MICS datasets, at http://www.childinfo.org
- World Bank compilation of 54 development indicators, at http://ddp-ext.worldbank.org/ext/DDPQQ/member.do?method=getMembers&userid=1&queryId=135
- WHO Statistical Information System online database of more than 100 indicators, at http://www.who.int/whosis/en/
- UNICEF series of global databases with many indicators used for Millennium Development Goals (MDG) and other global monitoring efforts, at http://www. childinfo.org
- WHO and UNICEF compiled data on immunization coverage, at http://www.who.int/immunization_monitoring/en/globalsummary/timeseries/tswucoveragedtp3.htm
- Penn World Tables data on trends in gross domestic product (GDP), at http://pwt.econ.upenn.edu/php_site/pwt_index.php

X.2 Accounting for Environmental Factors

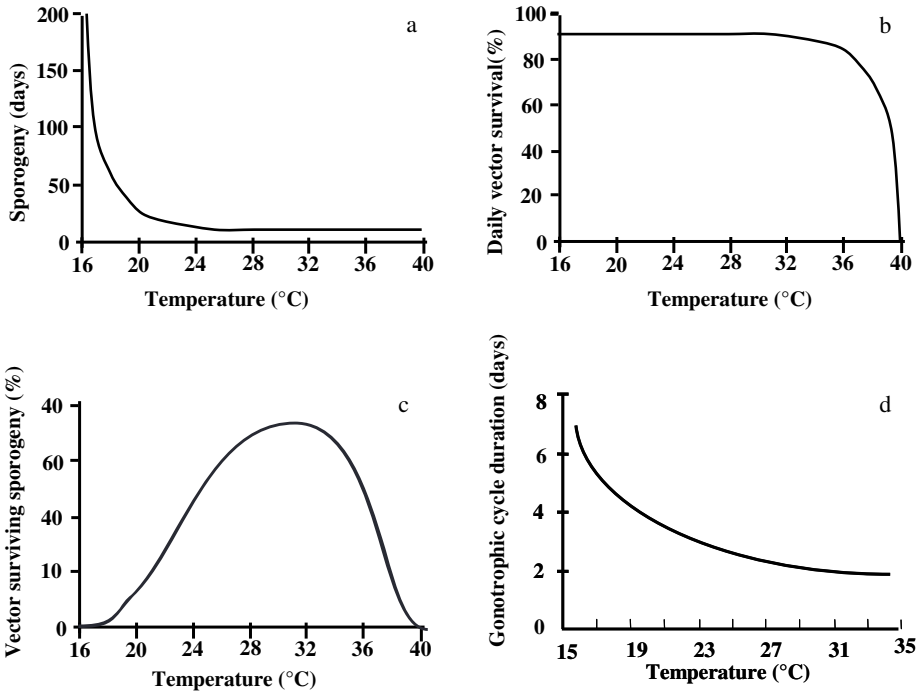
X.2.1 Temperature

Ambient temperature plays a major role in the malaria vector life cycle, which affects transmission risk. The parasite development in the mosquito host (sporogonic cycle) depends on temperature. The time delay from infection to infectious stages of the mosquito is estimated at 9–10 days at a temperature of 28 °C; development stops at temperatures below 16 °C.[2,3] Figure X.1a shows the rapid decrease of the sporogonic cycle length with increase in temperature. At 28 °C, the length of the cycle is shorter.[4]

Daily vector survival (90%) is stable at temperatures between 16 and 36 °C. Higher temperature increases mosquito mortality (Figure X.1b). The number of mosquitoes that survive the sporogonic cycle has a bell-shaped relationship with temperature (Figure X.1c). At low temperatures, the vector may not survive long enough for completion of the sporogonic cycle, but at high temperatures, mosquito mortality increases. The highest proportion of vectors that survive the incubation period is

observed at temperatures between 28 and 32 °C.[5] The gonotrophic cycle, which is the vector feeding interval, is shortened by high temperatures (Figure X.1d). Because high temperatures increase the digestion rate of blood meals, high temperatures result in more frequent vector-host contacts.

FIGURE X.1: EFFECT OF AMBIENT TEMPERATURE ON MALARIA VECTOR ADULT STAGE: (A) DURATION OF THE SPOROGONIC CYCLE IN DAYS; (B) DAILY MOSQUITO SURVIVAL; (C) PERCENTAGE OF VECTORS THAT SURVIVE SPOROGENY; (D) GONOTROPHIC CYCLE



Source: Ye, et al. 2008; Macdonald 1957; Detinova 1962; Martens 1997; Detinova 1962

Immature stages of the vector are equally temperature dependent. The duration of the larval stage is long at 16 °C and consistently decreases with an increase in temperature. The shortest time is observed at 36 °C. Detinova expresses the relationship this way:

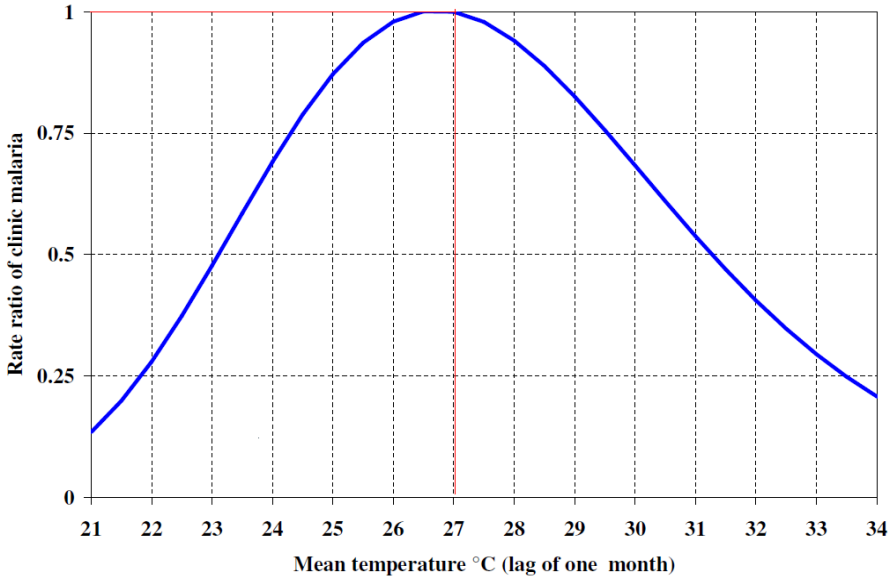
$$n = \frac{111}{T\,°C - 18\,°C},$$

In this equation, n is the duration in days of the larval stage, $T\,°C$ the ambient temperature, and $18\,°C$ the threshold below which larval development stops; 111 is the thermal sum in degree-days.[6] Temperature is, therefore, the main determinant of the basic reproduction rate (R_0). If R_0 less than 1, malaria will die out; if R_0 >1 malaria will be sustained.[7]

Temperature also can be used to assess risk of clinical malaria (presence of parasite and clinical signs). Ye, et al., 2007 used logistic regression with fractional polynomial transformation to show that the relationship between mean temperature (with a 1one-month lag) and clinical malaria

incidence is a bell-shape curve.[8] The risk of clinical malaria increased with an increase in mean temperature up to 27 °C. At 23 °C, the risk for clinical malaria was 53% of the risk at 27 °C. The risk was, therefore, least at the lower and higher extremes of the temperature range (Figure X.2).

FIGURE X.2: EFFECT OF MEAN TEMPERATURE ON CLINICAL MALARIA RISK AMONG STUDY CHILDREN IN NORTHWEST BURKINA FASO, 2007



Note: Horizontal and vertical red lines indicate the reference point (rate ratio = 1; T = 27 °C).

Source: Ye, et al. 2007

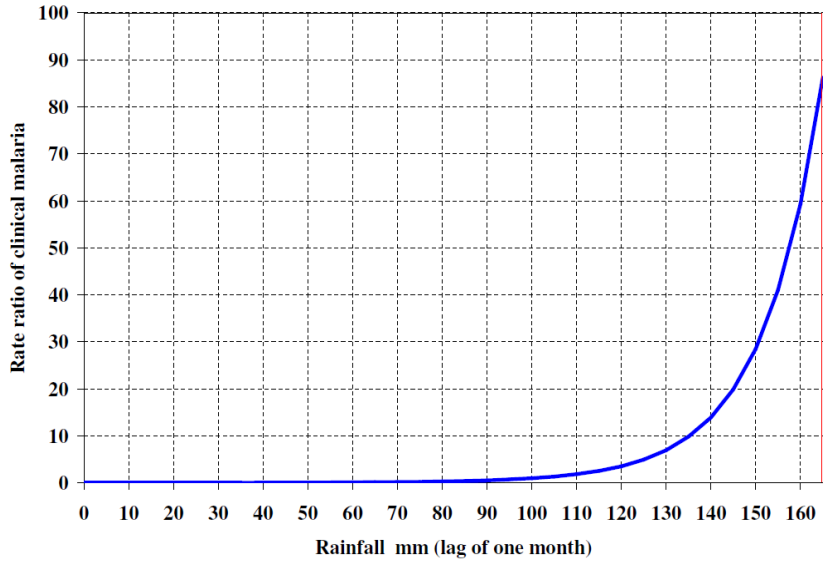
An evaluation of malaria interventions needs monthly data on temperature (minima, means, and maxima) for at least the period being evaluated, and, if possible, five years before the start year of the period being evaluated. National-level data may not be informative; for example, in most countries in Sub Saharan Africa, national-level temperatures will be within the range for sustained transmission. Subnational-level data can provide more insight on changes in subnational malaria transmission. Time-series data on temperature are available from national meteorological offices in most countries; researchers can supplement them with air temperature from satellites, such as Moderate Resolution Imaging Spectroradiometer (MODIS), at http://modis.gsfc.nasa.gov/ or the National Oceanographic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA-AVHRR), at http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html.

X.2.2 Rainfall and Humidity

Most malaria vectors depend on rainfall because it provides favorable breeding sites for mosquitoes to lay eggs. Rainfall also ensures suitable relative humidity of at least 50 to 60% for mosquitoes to survive. A relative humidity value below 60% shortens the mosquito lifespan. Too much rain can destroy breeding sites or flush away eggs or larvae.[9]

The onset of the rainy season historically correlates with an increase in vector abundance.[10,11] In Nairobi, Kenya, outbreaks of malaria occurred in 1940 after heavy rains.[12] In the Ugandan highlands, rainfall anomalies (based on a difference from the mean) that resulted from the 1997 El Niño were positively correlated with vector density one month later, which may have initiated the resulting epidemic.[13] The association between rainfall and malaria was observed consistently in several other studies.[14,15,16] The total amount of rainfall is also highly correlated with malaria risk; however, depending on the country or location, a minimum amount of rainfall is required to influence malaria transmission risk. Ye, et al. found that the effect of rainfall (with a 1-month lag) on clinical malaria risk was observed only after a monthly rainfall of above 100 millimeters.[16] Above this threshold, clinical malaria risk increased significantly for each increase of 10 millimeters of rain (the rate ratio at 150 millimeters was 0.3, compared with 0.7 at 163 millimeters). Below 100 millimeters, the risk was close to 0 (Figure X.3). Nonetheless, in some studies, the rainfall-malaria relationship was not established. Lindsay, et al. found that after one El Niño event in Tanzania's highlands, with 2.4 times more rainfall than normal, fewer malaria cases were reported than the year before.[17]

FIGURE X.3: EFFECT OF TOTAL RAINFALL ON CLINICAL MALARIA RISK AMONG STUDY CHILDREN IN NORTHWEST BURKINA FASO, 2007



Note: Horizontal and vertical red lines indicate the reference point (rate ratio = 1, Pmm = 163 millimeters).

Source: Ye, et al. 2007

Because rainfall affects mosquito breeding rates, it is important to consider this factor when assessing malaria transmission risk at the subnational level. For the period that covers the evaluation, it is ideal to have subnational-level data on monthly rainfall and the number of days of rain each month. Time series data on rainfall is available from national meteorological offices in most countries. Researchers can supplement this data with satellite data from NOAA-AVHRR.

X.2.3 Land Cover

Land cover, in particular vegetation, has been associated with malaria transmission because it provides suitable habitat for mosquito breeding. The presence of vegetation creates a microclimate of moderate temperature and humidity that is suitable for mosquitoes.[18] Land cover is measured either directly by field observation or indirectly by satellite imagery. Satellite imagery quantifies the amount of vegetation using the Normative Difference Vegetation Index (NDVI). In Kenya, NDVI was found to correlate with human-biting rate[19] and the annual number of malaria cases.[20] Similarly, Eisele, et al. demonstrated that the number of potential *Anopheles* larval habitats increased with increasing mean value of NDVI, which is a good proxy for assessing malaria transmission risk;[21] however, field validation is required for accurate estimates and might not be effective at the national level.

X.3 Data Analysis

Evaluators should analyze data on contextual factors using standard statistical methods to examine continuous and categorical variables, as appropriate. Specific guidance for evaluating these indicators from DHS, MICS, and MIS appears in the survey documentation. An assessment of temporal changes of all relevant contextual factors for the same period that malaria control interventions are scaled up should follow a univariate analysis.[22,23,24] Evaluators should consider including statistically significant temporal changes of contextual factors in multivariable analyses to assess potential confounding and effect modification.

Table X.2 shows an example of the analyses conducted during an impact evaluation in Tanzania.[25] Household factors are important determinants of child health and malaria risk, so evaluators assessed these factors for change over time from the baseline (1999) to the end line (2010) of the evaluation. A simple table presents the results as part of a larger discussion on contextual factors.

TABLE X.2: HOUSEHOLD ATTRIBUTES AND ASSET OWNERSHIP, MAINLAND TANZANIA, 1999–2010

Survey Year	1999			2010					
	%	95% CI	N	%	95% CI	n	% change	Dir. change	Sig.
Improved water source, (% households)	65.8	(59.1-71.8)	3526	56.9	(53.3-60.4)	9377	-13.5	↓	ns
Time to water source less than 15 minutes, (% households)	34.3	(29.6-39.4)	3526	36.5	(33.8-39.4)	9377	6.4	↑	ns
Improved toilet facilities (not shared), (% households)	1.5	(1.0-2.3)	3526	12.1	(10.5-14.0)	9377	7.1	↑	*
Improved roof (not thatch/grass/mud), (% households) §	50.6	(47.3-53.8)	9483	61.9	(58.9-64.9)	9377	22.3	↑	*
Modern floor material (not earth, sand, or dung), (% households)	20.9	(17.0-25.4)	3526	31.8	(28.7-35.0)	9377	52.2	↑	*

Survey Year	1999			2010					
	%	95% CI	N	%	95% CI	n	% change	Dir. change	Sig.
Electricity, (% households)	7.7	(5.5-10.7)	3526	14.2	(11.9-17.0)	9377	84.4	↑	*
Telephone, (% households) §	8.9	(7.6-10.5)	9483	45.5	(42.9-48.0)	9377	411.2	↑	*
Often/always had problems satisfying food needs in past year (% households)	22.6	(21.2-24.1)	9483	23.3	(21.6-25.1)	9377	3.1	↑	ns

Notes: Improved water source is protected, borehole, piped; § signifies 2004/5 DHS source.

Dir. Change: direction of change. Sig: Statistical significance. Statistics with non-overlapping 95% confidence intervals are considered significantly different change. NS denotes no statistically significant change; * denotes statistically significant change.

---

[1] Victora C, et al. Context matters: Interpreting impact findings in child survival evaluations. *Health Policy Plan*. 2005. 20 (suppl 1):i18–i31. doi:10.1093/heapol/czi050.

[2] Macdonald WW. Aëdes aegypti in Malaya. III. Experiments on control and eradication. *Ann Trop Med Parasitol*. 1957. 51(2):175–86.

[3] Bradley TJ. Physiology of osmoregulation in mosquitoes. *Annu Rev Entomol*. 1987. 32:439–62.

[4] Lindsay SW, and Birley MH. Climate change and malaria transmission. *Ann Trop Med Parasitol*. 1996. 90(6):573–88.

[5] Craig MH, Snow RW, and le Sueur D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today*. 1999. 15(3):105–11.

[6] Detinova TS. Age-grouping methods in Diptera of medical importance with special reference to some vectors of malaria. *Monogr Ser WHO N° 47*. 1962. Geneva.

[7] Macdonald G. Appendix I. Mathematical statement. In *The Epidemiology and Control of Malaria*. 1957. pp 201, Oxford University Press; London.

[8] Yé Y, Louis VR, Simboro S, and Sauerborn R. Effect of meteorological factors on clinical malaria risk among children: An assessment using village-based meteorological stations and community-based parasitological survey. *BMC Public Health*. 2007. 7:101.

[9] Craig MH, Snow RW, and le Sueur D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today*. 1999. 15(3):105–11.

[10] Gill CA. The relationship between malaria and rainfall. *Indian Journal of Medical Research*. 1920. 37:618–632.

[11] Haddow AJ. The mosquito fauna and climate of native hats at Kisumu, Kenya. *Bulletin of Entomological Research*. 1942. 33:91–142.

[12] Lindsay SW and Martens WJ. Malaria in the African highlands: past, present and future. *Bull World Health Organ*. 1998. 76:33–45.

[13] Lindblade KA, Walker ED, Onapa AW, Katungu J, and Wilson ML. Highland malaria in Uganda: Prospective analysis of an epidemic associated with El Nino. *Trans R Soc Trop Med Hyg*. 1999. 93:480–7.

[14] Githeko AK, and Ndegwa W. Predicting malaria epidemics in Kenyan highlands using climate data: A tool for decision makers. *Glob Change Hum Health*. 2001. 2:54–63.

[15] Zhou G, Minakawa N, Githeko AK, and Yan G. Climate variability and malaria epidemics in the highlands of East Africa. *Trends Parasitol*. 2005. 21(2):54–6

[16] Ye Y, Louis V, Simboro S, and Sauerborn R. Assessing meteorological effects *Plasmodium falciparum* malaria infections among children, using village-based meteorological stations and community-based active case detection methods. *BMC Public Health*. 2007. 7(1):101.

[17] Lindsay SW, Bødker R, Malima R, Msangeni HA, and Kisinza W. Effect of 1997–98 El Niño on highland malaria in Tanzania. *Lancet*. 2000). 355(9208):989–90.

[18] Clements AN. *The Biology of Mosquitoes–Volume 2 Sensory Reception and Behaviour*. 1999. CABI Publishing, Wallingford.

[19] Patz JA, Strzepek K, Lele S, Hedden M, Greene S, Noden B, Hay SI, Kalkstein L, and Beier JC. Predicting key malaria transmission factors, biting and entomological inoculation rates, using modeled soil moisture in Kenya. *Trop Med Int Health*. 1998. 3(10):818–27

[20] Hay SI, Snow RW, and Rogers DJ. Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. *Trans R Soc Trop Med Hyg*. 1998. 92(1):12–20.

[21] Eisele TP, Keating J, Swalm C, Mbogo CM, Githeko AK, Regens JL, Githure JI, Andrews L, and Beier JC. Linking field-based ecological data with remotely sensed data using a geographic information system in two malaria endemic urban areas of Kenya. *Malar J*. 2003. 2(1):44.

[22] Bryce J, Gilroy K, Jones G, Hazel E, Black RE, and Victora CG. The Accelerated Child Survival and Development programme in West Africa: A retrospective evaluation. *Lancet*. 2010. 375(9714):572–582. doi:10.1016/S0140-6736(09)62060–2.

[23] Rowe AK, Onikpo F, Lama M, Osterholt DM, and Deming MS. Impact of a malaria-control project in Benin that included the integrated management of childhood illness strategy. *Am J Public Health*. 2011. 101(12):2333–41.

[24] Rowe AK, Kachur SP, Yoon SS, Lynch M, Slutsker L, et al. Caution is required when using health facility-based data to evaluate the health impact of malaria control efforts in Africa. *Malar J*. 2009. 8:209.

[25] Tanzania Malaria Impact Evaluation Group. Evaluation of the Impact of Malaria Control Interventions on All-Cause Mortality in Children Under-five in Mainland Tanzania. 2012.

XI. Data Synthesis, Triangulation, and Interpretation

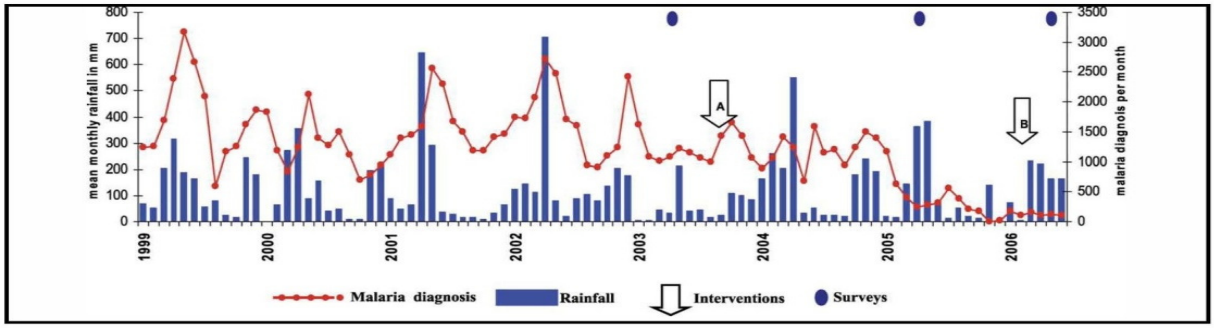
Key Points

- Assessing the overall impact of a country's scale-up of its full-coverage, malaria control program requires an ecological study design.
- Work by Bhattarai and colleagues (2007) provides an excellent example of plotting data from multiple sources, including cases from the Health Management Information System, timing of surveys, intervention start dates, and rainfall.
- Following an ecological analysis of all available datasets to assess trends in malaria morbidity and mortality over time, MERG suggests conducting sub-analyses to help validate any positive effects found by the plausibility design.

XI.1 Trends Over Time

An ecological study design is proposed to assess the overall impact of a country's scale-up of its full-coverage malaria control program. The design assesses simultaneous changes in intervention coverage and the malaria burden outcomes of morbidity and mortality at the population level. An example from Zanzibar (Figure XI.1) shows data plotted from multiple sources, such as cases from the Health Management Information System (HMIS), timing of surveys and intervention start dates, and rainfall (a contextual factor).[1] The outcome in the example from Zanzibar is monthly incidence cases, while explanatory variables include year, the year of program intervention, and monthly rainfall. Inclusion of the year of program intervention and monthly rainfall depend on data availability. Appendix B discusses trends in these factors and trends in all-cause child mortality (ACCM), yearly *P. falciparum* parasite rate (*Pf*PR), and inpatient malaria deaths. Evaluators also could assess temporal trends using a regression model, such as segmented regression, depending on data availability. This type of assessment allows for control of potential confounding factors, which improves the validity of the results.

FIGURE XI.1: MALARIA INTERVENTIONS, CROSS-SECTIONAL SURVEYS, MONTHLY RAINFALL, AND REPORTED CLINICAL MALARIA DIAGNOSES IN CHILDREN UNDER AGE 5 YEARS IN NORTH A DISTRICT, ZANZIBAR (1999–2006)



Source: Bhattarai, et al., 2007

Malaria control programs usually have a larger impact in areas of higher malaria transmission and among high-risk subpopulations, such as children and pregnant women. Therefore, evaluators should stratify data by malaria transmission risk, including urban and rural residence. They could determine transmission strata using intervention coverage, data reporting, risk maps using mean *Pf*PR (see Information for Malaria (INFORM) and Malaria Atlas Project, http://www.map.ox.ac.uk/), and existing risk maps based on climatic and environmental suitability. Evaluators should base stratification on pre-intervention transmission levels.

To confirm that malaria control programs reach the most vulnerable populations that typically are at greatest risk of malaria, evaluators should stratify coverage indicators measured from cross-sectional household surveys by household wealth quintiles and mother's education level. This allows programs to use the information to target vulnerable populations that are the hardest to reach. Program should define equity in coverage as equality across household wealth quintiles and mother's education or favoring lower socioeconomic quintiles and less educated households. A recent publication discusses issues in measuring household wealth based on household surveys and their implications for program coverage.[2]

XI.2 Further Analysis

The validity of the plausibility study design to measure the impact of scaling up a malaria control program depends on empirical proof that the interventions assessed have an impact on the outcomes measured. For example, insecticide-treated nets (ITNs) are proven effective in reducing ACCM in randomized controlled trials.[3] After an ecological analysis of all available datasets to assess trends in malaria morbidity and mortality over time, as outlined above and in Appendix B, a separate analysis can help evaluators assess the effectiveness of proven interventions under program conditions. Where effectiveness is confirmed, especially when consistent with trial data, subanalyses will bolster the results obtained with the plausibility study design. National-level analyses are ideal, but subnational studies also can support the plausibility argument.

Examples of further analyses that can be tailored to available data in a specific country include:

- District-level analysis of factors associated with the incidence of outpatient malaria cases and inpatient malaria deaths measured through HMIS
- Analysis of cross-sectional datasets to assess the association of ITN exposure and other malaria control measures with malaria health outcomes using exact matching and propensity score matching
- Survival analysis of survey datasets with complete birth histories to assess the effect of ITN exposure and other malaria control measures and child survival

XI.3 Multicountry Meta-analysis

After evaluations have been completed in several countries and comparable datasets are available, a meta-analysis may provide more insight into malaria program effectiveness, especially where in-country analyses lack sufficient statistical power to detect program effects. Meta-analyses can consist of analysis of aggregated point estimates from national-level or subnational studies, typically presented as forest plots. They also may consist of pooled individual or household-level data, depending on data availability and study objectives. An example of a meta-analysis of country-level outcomes by Lim and colleagues assessed the associations of ITN exposure to reductions in ACCM and parasite prevalence across 29 national survey datasets.[4] Evaluators may repeat this analysis after more national survey data are available. The meta-analysis by Eisele and colleagues is an example of a pooled individual-level meta-analysis.[5] Researchers assessed the association of intermittent preventive treatment in pregnancy with sulfadoxine pyrimethamine and ITNs, or one or the other, with reductions in low birth weight and neonatal mortality. In this analysis, "country" was included as a random effect to replicate a meta-analysis of country-level estimates. Both the Lim and Eisele studies used exact matching to mitigate confounding bias where children and mothers exposed to malaria control interventions likely have better health outcomes.

XI.4 Synthesis of Results

As outlined in previous sections, an evaluation will produce a large body of results over the course of the malaria control program lifespan. The strength of evidence for causality will range, based on the rigor of the method employed. All available data points from an evaluation are considered to establish an overall direction and magnitude of the impact of a malaria control program on key impact indicators. Table XI.1 lists considerations for assessing whether observed associations suggest causal relationships. The framework established by Bradford Hill and interpreted by Hofler[6] can help systematically incorporate all available data points into a comprehensive image of the plausible impact that a national malaria control program had on the established impact indicators.[7] Each of the nine considerations shown in Table XI.1 can be matched with results from national impact evaluations, and, where satisfied, will add to the evidence of the overall impact of a malaria control program.

All results that demonstrate a positive association between malaria control program inputs and coverage and impact indicators will establish strength of association. Meta-analyses that demonstrate similar results across evaluation methods over time and across countries meet considerations for consistency and analogy. Disaggregated results that show larger impacts on subgroups expected to benefit most from malaria control, such as those in areas of higher transmission, will establish specificity. Observed changes in outcome and impact indicators that occur after the scale-up of malaria control activities will establish temporality. Results of the district-level, dose-response analysis to assess the association of program intensity with outcome and impact indicators will provide robust evidence of gradient. Although previous randomized controlled trials may establish experimental evidence, secondary analyses, such as matched regression analysis, also will contribute to experimental evidence of an effect. Robust existing empirical evidence of the efficacy of malaria control interventions against morbidity and mortality primarily establish plausibility and coherence.

TABLE XI.1: NINE CONSIDERATIONS TO ASSESS WHETHER OBSERVED ASSOCIATIONS INVOLVE CAUSAL RELATIONSHIPS

1. Strength of association	Strong associations are more likely to have causal components than weaker associations.
2. Consistency	Observing similar evaluation results across evaluation methods, over time, and across countries from meta-analyses increases the likelihood of causal relationships.
3. Specificity	Observing an association specific to outcomes of interest among specific groups increases the argument for causal effect.
4. Temporality	Changes in program must precede changes in disease or coverage outcomes.
5. Gradient	Changes in disease or coverage outcomes increase the same amount for increases to program exposure or intensity.
6. Plausibility	Biological plausibility links exposure to intervention with health outcome.
7. Coherence	Causal inference is possible only if the literature or substantive knowledge supports this conclusion
8. Experiment	Causation is a valid conclusion if researchers have seen observed associations in prior experimental studies.
9. Analogy	For similar programs operating, similar results can be expected to bolster the causal inference concluded.

Source: Bradford Hill Considerations for Causality: a counterfactual perspective

[1] Bhattarai A, Ali AS, et al. Impact of artemisinin-based combination therapy and insecticide-treated nets on malaria burden in Zanzibar. *PLoS Med.* 2007. 4(11):e309.

[2] Barros AJD, and Victora CG. Measuring coverage in MNCH: Determining and interpreting inequalities in coverage of maternal, newborn, and child health interventions. *PLoS Med.* 2013. 10(5):e1001390.

[3] Lengeler, C. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev.* 2004. (2):CD000363.

[4] Lim SS, Fullman N, Stokes A, Ravishankar N, Masiye F, et al. Net benefits: A multicountry analysis of observational data examining associations between insecticide-treated mosquito nets and health outcomes. *PLoS Med.* 2011. 8(9):e1001091.

[5] Eisele TP, Larsen DA, Anglewicz PA, et al. Malaria prevention in pregnancy, birthweight, and neonatal mortality: a meta-analysis of 32 national cross-sectional datasets in Africa. *Lancet Infect Dis.* 2012. 12(12):942–949.

[6] Hofler M. The Bradford Hill considerations on causality: A counterfactual perspective. *Emerging Themes in Epidemiology.* 2005. 2:11.

[7] Hill, AB. The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine.* 1965. 58(5): 295–300.

Appendices

A. Limitations and Assumptions of the Plausibility Study Design

B. What Results Might Look Like: Example Scenarios

C. Malaria Risk Parameters

D. Details of Analyses to Help Validate and Quantify Results Obtained from the Plausibility Study Design

E. Country Case Studies on the Analytic Approaches

F. Case Study 1: Bioko Island

G. Case Study 2: Zambia

H. Case Study 3: Zanzibar, North A District

I. Case Study 4: Mainland Tanzania

J. Case Study 5: Malawi

Appendix A. Limitations and Assumptions of the Plausibility Study Design

Limitation	Assumption	Potential means of mitigation
Causal inference of the malaria control program on observed morbidity and mortality outcomes cannot be established.	Changes in outcomes would not have happened without the scale-up of the malaria control program.	There is strong empirical evidence that interventions all have an impact on the malaria and morbidity outcomes. Interpretation of results with the inclusion of extraneous factors might also influence malaria impact indicators. Sub-analyses assessing specific components of the intervention on malaria outcomes can bolster the plausibility argument.
Most of the primary malaria control trials established estimates of efficacy, which do not always translate to the same effectiveness under program conditions.	The efficacies established by the trials approximate the effectiveness of the interventions under program conditions.	Sub-analyses can quantify the effectiveness of interventions under program conditions (such as multi-level regression analyses of cross-sectional data to assess the association of insecticide-treated nets with malaria parasite prevalence after accounting for selection bias).
Quantifying contribution of different components of the national malaria control program on outcomes will be limited.	Impact evaluation captures the effects of all in-country malaria control efforts.	Potential for a dose-response or comparative analysis between regions, given an uneven rollout over time. Use of models (such as the Lives Saved Tool [LiST]) may help quantify the relative contribution of different interventions on modeled reductions in all-cause child mortality or malaria-specific mortality.
Measuring external contextual factors—such as rainfall, socioeconomic status, and policy changes—will be challenging.	It will be possible to measure changes in contextual factors over time.	The evaluation will involve many stakeholders across government agencies and organizations, thus access to longitudinal data at the district level should be possible.
Survey data may be only from two data points and may not reflect a recent rapid scale-up of intervention and its impact; surveys may not have sufficient sample size to allow regional or district-level trend analysis.	Two data points adequately represent pre- and post- intervention situation in a country.	Two data points will be interpreted along with trends from estimates derived from health information management systems (HMIS). Surveys have sampling error, but can be statistically presented. Birth history and regression analysis can produce estimates for other years.
HMIS-derived estimates of malaria incidence and mortality are biased.	Although HMIS-derived estimates produce underestimates of these indicators, the estimates can help assess relative changes over time, assuming access and utilization (underestimation) are, for the most part, constant over time, except where diagnostics is introduced along with changes in case definition from clinical to laboratory confirmed.	Trends in HMIS-derived estimates will be interpreted, along with trends in health service access and utilization. The estimates can assess not only trends in absolute case and death numbers, but also the proportion that malaria (and in stable endemic Sub Saharan Africa, also anemia) makes up out of all-cause outpatient visits, hospital admissions, and hospital deaths (Aregawi, et al., 2011). Multiple data sources will be used to measure outcomes, including surveys, HMIS, and studies. Analysis should carefully consider the case definition, focusing on confirmed malaria cases where possible.

Appendix B. What Results Might Look Like: Example Scenarios

Figure B.1 illustrates what three potential results of a plausibility evaluation study design using multiple data points might look like after successful scale-up of the national malaria control program.

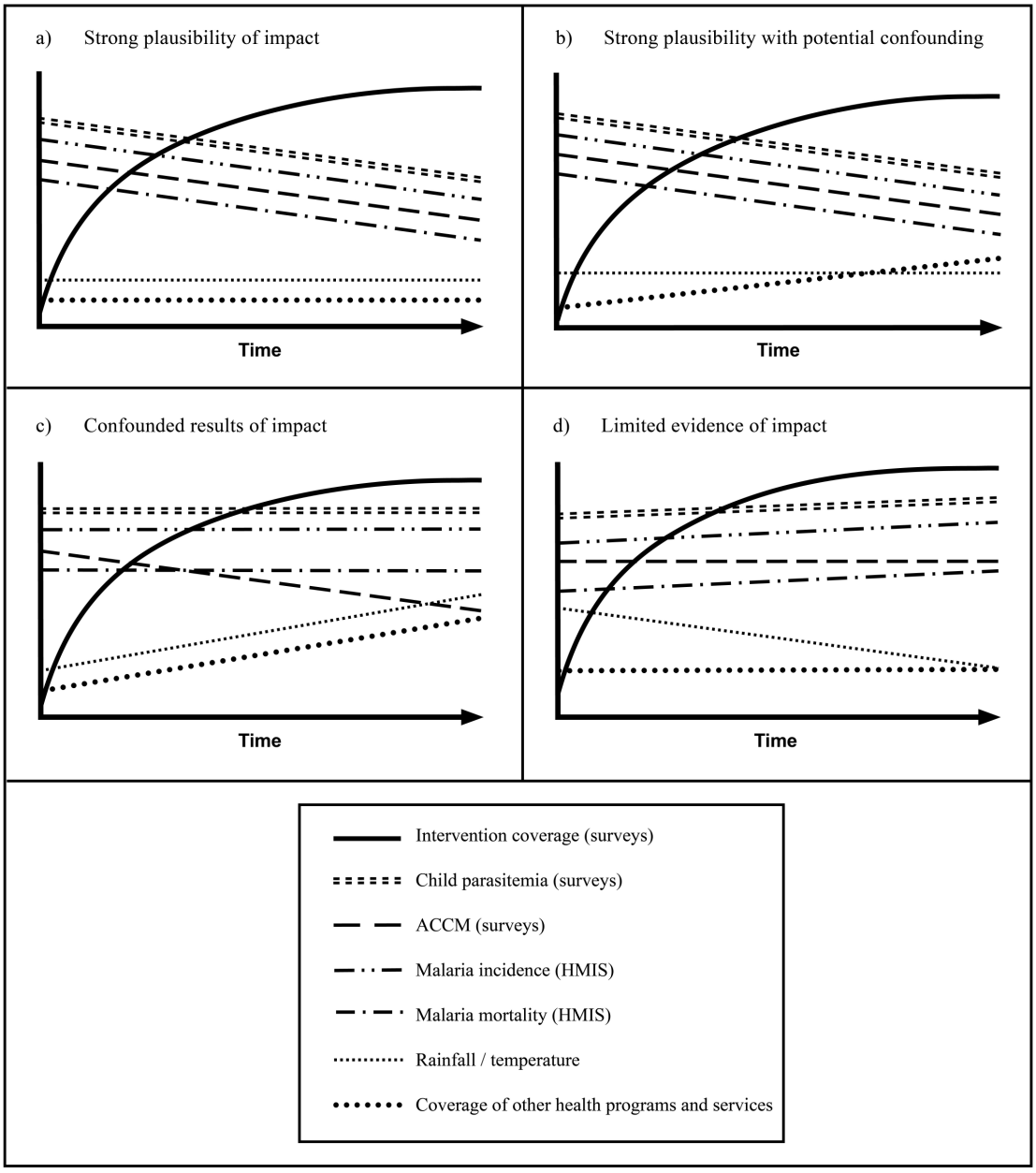
Figure B.1a shows the hypothetical results from such an analysis for the scenario of strong plausibility of impact on malaria-related morbidity and mortality outcomes. In this scenario, there is sufficient evidence that the national program was scaled up successfully (i.e., statistically significant increases in coverage). Further, all key impact indicators for malaria have declined. Lastly, extraneous factors that might have contributed to declining malaria levels do not appear to have influenced such declines (they remain constant in this example, and multi-level analyses also may control for them). Thus, under this scenario, it would be very plausible to conclude that the scale-up of the national malaria control program had a demonstrable impact on malaria morbidity and mortality for the following reasons: 1) There is strong empirical evidence linking coverage of malaria control interventions to reductions in malaria morbidity and mortality, and population coverage increased to desired levels; 2) changes in multiple impact indicators for malaria morbidity and mortality consistently changed in desired directions from pre-program scale-up to post-program scale-up; and 3) results show it is unlikely such changes would have been due to extraneous factors.

Figure B.1b also shows the scenario of strong plausibility of impact on malaria-related morbidity and mortality outcomes, although the results for decreases in all-cause child mortality (ACCM) are likely a result of improved access to several child survival services and interventions (such as immunization). In this scenario, it would remain plausible to conclude that the scale-up of the national malaria control program had a demonstrable impact on malaria morbidity and mortality as noted above. However, the conclusion should have the caveat that a multitude of factors caused declines in ACCM, one being improved access of malaria control. The latter scenario may be the most realistic.

Figure B.1c shows the scenario of limited plausibility of impact on malaria-related morbidity and mortality outcomes, as measures of malaria morbidity and mortality did not decline despite increased coverage of malaria control interventions. This scenario would require further analysis to try and assess the counterfactual of what might have happened had malaria control not been scaled up (that is, evaluators may deem the program effective if they estimate that the malaria burden would have worsened if not for program scale-up).

Figure B.1d shows the scenario of no plausible evidence of impact on malaria-related morbidity and mortality outcomes, as measures of malaria morbidity and mortality worsened despite increased coverage of malaria control interventions, and despite favorable extraneous factors.

FIGURE B.1 A–D: ASSESSMENT OF MULTIPLE DATA POINT FOR INDICATORS FOR MALARIA CONTROL POPULATION



Appendix C. Malaria Risk Parameters

Parameter	Methods	Formula
Human biting rate (HBR)	***a***, mosquito feeding frequency ***m***, human blood index, proportion of mosquito blood meals obtained from humans	ma
Vectorial capacity (VC)	***ma*** (HBR) ***p***, daily mosquito population survival ***n***, incubation period of parasite in vector	$\frac{ma^2 p^n}{-\ln p}$
Basic reproductive rate (R_0)	***b***, % of vector developing parasite following ingestion of gametes ***r***, recovery rate from infection	$\frac{ma^2 bp^n}{-r \ln p}$
Entomological inoculation rate (EIR)	***ma*** (HBR) ***s***, proportion of infected mosquitoes (sporozoite rate)	mas
Parasite prevalence	Confirmed with malaria diagnostic test (microscopy or rapid diagnostic test) out of the total tested	
Incidence of malaria morbidity	Confirmed with malaria diagnostic test and clinical examination	

Appendix D. Details of Analyses to Help Validate and Quantify Results Obtained from the Plausibility Study Design

As discussed, an ecological study design will be used to assess the overall impact of a country's scale-up of their full coverage malaria control program. Such a design will assess the simultaneous changes in intervention coverage and the malaria burden outcomes (morbidity and mortality) at the population level. Evaluators could analyze such data with a simple multiple regression model. As noted, the outcome in the example from Zanzibar is monthly incidence cases, whereas the explanatory variables are year, year of program intervention, and monthly rainfall (inclusion would depend on data availability). Evaluators also could assess trends in all-cause, under-5 mortality, yearly *P. falciparum* parasite rate (PfPR) (predicated or measured), and inpatient malaria deaths using this technique, depending on data availability.

Depending on the distribution of the outcome, evaluators could use Linear, Poisson, or Negative Binomial Regression. In a model of incidence cases that decrease over time, one would expect a negative and significant (in reference to baseline year) regression coefficient for calendar year, after controlling for rainfall. This would signify a decreasing trend in cases over time, generally associated with increasing program intensity. The following results would provide even stronger evidence of a program effect: a) no statistical decline in the outcome during the years preceding malaria control scale-up (1999–2003 in the Bhattarai example); and b) a statistically significant decline in the outcome for the years following the malaria control scale-up (2004–2007), after controlling for monthly rainfall and any other possible contextual factors.

Secondary analysis of program, survey, and contextual data

As noted, the validity of the plausibility study design for assessing the impact of malaria control program scale-up depends on empirical proof that the interventions assessed have an impact on the outcomes measured. For example, insecticide-treated bed nets (ITNs) have established protective efficacy in reducing all-cause child mortality (ACCM) from randomized controlled trials.[1] Following an ecological analysis outlined above, sub-analyses will help validate any positive effects found by the plausibility design. Doing so will help assess the effects of the interventions under program conditions. Confirmed effects—especially when consistent with trial data—will bolster the results obtained with the plausibility study design. Overviews of these analyses appear below, though evaluators should tailor the details of the analytic plans to the data available within each country.

District-level analysis of factors associated with outpatient malaria case incidence and inpatient malaria deaths measured through the Health Management Information System (HMIS): Each country should perform a district-level analysis to assess the association of malaria program intensity and monthly, HMIS-derived, outpatient malaria-case incidence and inpatient malaria deaths (standardized by population size), while accounting for contextual factors such as rainfall,

temperature, treatment-seeking behaviors, and HMIS-reporting quality. Evaluators could conduct the analysis at the level of discrete, district-time units.

The simplest approach involves using program data to obtain monthly or yearly estimates of malaria program intensity, as was previously done by Graves and colleagues in Eritrea.[2] For example, ITNs distributed per capita or houses sprayed per capita for each district per month (or year) could be used as proxy measures for malaria program intensity. Alternatively, a Bayesian geostatistical approach could be used to estimate district-level coverage of ITN and indoor residual spraying (IRS), based on nationally representative household surveys and commodity-distribution data, as has been done previously.[3] Evaluators could obtain measures of district-level rainfall and temperature (monthly, quarterly, or yearly) from national meteorological services, or derive them from remote sensing data (such as enhanced vegetation index).[4]

It also is important for evaluators to account for the effect of treatment-seeking behavior and HMIS quality at the district level. Predicted estimates of treatment-seeking behavior for childhood fevers measured from nationally representative household surveys using a Bayesian approach can provide a proxy for treatment-seeking behavior for each district. Estimates of HMIS data quality and completeness for each district can be obtained from the National Malaria Control Centre or the Ministry of Health (MoH). The final analysis of the district-time units may require analysis in a Bayesian framework to properly account for spatial and temporal autocorrelation across district-time units. If successful, this analysis should also allow the assessment of the relative contribution of different malaria interventions (such as IRS, as compared with ITNs) above and beyond ecological factors related to transmission risk (such as elevation, climate, urban-rural differences, and access issues). Appendix G presents a case study of such an analysis recently undertaken for Zambia.

Analysis of cross-sectional datasets to assess the association of ITN exposure (and other control measures) with malaria morbidity outcomes: This analysis entails using multiple regression modeling to assess the association of exposure to malaria programs (such as ITNs) with decreases in malaria health outcomes (such as the prevalence of parasite infections and severe anemia) using population-based, cross-sectional data. This analysis allows one to account for individual, household, and community-level contextual factors (potential confounders), such as age, sex, mother's education, household socioeconomic status, and geographic location. Evaluators should limit this analysis to rural households, because that is where transmission is concentrated and where ITNs are expected to have the largest impact (rural strata). Because exposure to malaria control interventions is not random, it is likely that those exposed differ from those unexposed on confounding factors that influence malaria health outcomes, thus resulting in selection bias. To mitigate the selection bias associated with such non-random exposure, evaluators should use techniques such as propensity score matching. Best practice for the analysis of data obtained from a two-stage cluster sampling design indicates that evaluators should empirically estimate standard errors to account for correlated data at the primary sampling unit level (or the primary sampling unit included in the model as a random effect).

Survival analysis of survey datasets with complete birth histories to assess the effect of ITN exposure and child survival (1–59-month ACCM rates): This analysis entails using a Cox Proportional Hazard model to assess the association of exposure to household ITN possession and 1–59-month child survival using population-based, cross-sectional data. Evaluators should create a longitudinal survival dataset from a full-birth history from a population-based nationally representative survey, such as the DHS. The analysis should include time-varying covariates for calendar year, child age, ITN household possession (year the net was acquired), and rainfall (or malaria transmission season). Other potential confounders to assess and include in the model are birth order, mother's education, household socioeconomic status, geographic location, migration (if available), and access to other child survival interventions (measured only at the time of the survey). Evaluators should conduct this analysis among the rural strata, as a significant effect of ITNs on ACCM is only expected within areas of stable malaria transmission. Where sufficient statistical power exists, evaluators should stratify the analysis by transmission level within a country (for example, high, medium, and low). To mitigate the bias associated with non-random exposure to ITNs, evaluators should employ a matching technique (such as propensity score matching). Best practice for the analysis of data obtained from a two-stage cluster sampling design indicates that evaluators should empirically estimate standard errors to account for correlated data at the primary sampling-unit level (or the primary sampling unit included in the model as a random effect). Notable assumptions for this analysis include: the ITN was available for use in the household from the procurement date to the survey date; children under age 5 years in household continuously lived in the sampled household; and important dates (such as net procurement date, date of birth, and date of death) were recalled accurately. Appendix J presents a case study of such an analysis recently undertaken in Malawi.

---

[1] Lengeler C. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev*. 2004. (2):CD000363.

[2] Graves PM, Osgood DE, Thomson MC et al. Effectiveness of malaria control during changing climate conditions in Eritrea, 1998–2003. *Trop Med Int Health*. 2008. 13(2):218–28.

[3] Noor AM, Alegana VA, Patil AP, and Snow RW. Predicting the unmet need for biologically targeted coverage of insecticide-treated nets in Kenya. *Am J Trop Med Hyg*. 2010. 83(4):854–860.

[4] Mabaso MLH, Craig M, Vounatsou P, and Smith T. Towards empirical description of malaria seasonality in southern Africa: The example of Zimbabwe. *Trop Med Int Health*. 2005. 10(9):909–918.

Appendix E. Country Case Studies on the Analytic Approaches

Zambia District Platform Analysis

Data from routine health management information systems (RHMIS) data are a potentially underused source for evaluating the effect of malaria control program intensity on the malaria morbidity burden. Since 2009, facilities in Zambia have reported both clinical and parasitologically confirmed positive malaria—by rapid diagnostic tests (RDTs) or microscopy—through the RHMIS on a monthly basis. This study used these data to evaluate the association between vector control coverage and monthly confirmed malaria cases at the district level in Zambia for the period 2009–2011.

The methods for this evaluation are complex; a detailed paper on the methods and results is forthcoming. The evaluators first used a Bayesian geo-statistical model to create smoothed estimates of ITN ownership from MIS data and to estimate differences in fever treatment-seeking behavior by district from 2009–2011. District-level programmatic data on the distribution of ITNs were incorporated to improve coverage estimates. Evaluators used a conditional autoregressive model of present RHMIS data values, a spatial neighboring relationship between facilities, and a first-order temporal autoregressive term to impute missing facility monthly RHMIS case incidence. The analysis included mean monthly rainfall and mean minimum and maximum temperature from remote sensing imagery to control for climate variability. The statistical model also included differences in reporting and testing rates by district and month. Finally, evaluators modeled the association between confirmed cases and vector control coverage a negative binomial regression. This analysis accounted for spatial and temporal autocorrelation between district-month units within a Bayesian framework computed with Integrated Nested Laplace Approximations.

The impressive results demonstrated that ITN program intensity was associated with decreases in monthly, standardized, confirmed malaria-case incidence in a dose-response relationship. Total malaria outpatient cases increased from 2.9 million 2009 to 4.1 million in 2010, and decreased to 3.7 million in 2011. Confirmed cases increased from 871,000 in 2009 to 1.2 million in 2010 and 1.8 million in 2011. After controlling for reporting, testing, climate, and district-level factors influencing treatment seeking, the evaluators estimated that an increase in district-level ITN coverage of 1 ITN per household was associated with a 29-36% reduction in population standardized confirmed case incidence, on average. This evaluation demonstrates that RHMIS data, if improved through comprehensive parasitologically confirmed case reporting, can become an important data source for evaluating associations between malaria program scale-up and spatial and temporal trends in disease burden.

Malawi Survival Analysis to Assess the Association of ITN Exposure and All-Cause Child Mortality

Malawi has made major progress in scaling up coverage of insecticide-treated bed nets (ITNs) over the past six years. Demonstrating the protective effectiveness of household ITN ownership for preventing all-cause child mortality (ACCM) under routine program conditions is an important step in the causal pathway towards assessing population-level impact. Evaluators used data from the 2010 DHS to assess whether household ownership of an ITN protected against ACCM in Malawi from 2007–2010.

Evaluators determined household ITN ownership retrospectively from household self-report on current ITN ownership, age of nets owned (in months up to 36 months), and time of retreatment, if any (in months ago). Exact matching on wealth (above or below median principal component analysis score), urban or rural, PfPR2-10 (greater than or less than 40%), DPT3 coverage at primary sampling unit (PSU) level (above/below median), and mother's education (secondary+, primary, none) reduced confounding. The evaluators then assessed the relationship between household ITN ownership and child mortality (1–59 months) over the 36 months preceding the survey with Cox proportional hazards models, with analysis time measured in months and matched strata included as a shared frailty. Other covariates controlled for in the Cox model included household wealth quintile, child age category, mother's age category, parity, PSU-level diarrhea prevalence (two weeks preceding survey), DPT3 coverage at PSU level (continuous), PfPR2-10 at PSU level (continuous), and season (high transmission season).

The resultant retrospective cohort included 29,492 children under age 5 years who provided 652,775 child-months of observation and among whom there were 821 deaths over the observation period. After controlling for confounders, children in households with an ITN were significantly less likely to die compared to those without one (Hazard ratio (HR): 0.75, 95% confidence interval [CI]: 0.62-0.90). In addition, ITNs reported as less than 1.5 years old provided greater protection than older nets (HR less than 1.5 years: 0.73, 95% CI: 0.60-0.88, over 1.5 years: 1.18, 95% CI: 0.70-1.97). These results demonstrate the impact of household ITN ownership on ACCM in this setting and suggest that ITN coverage in Malawi may have contributed significantly to decreased child mortality.

Appendix F. Case Study 1: Bioko Island (Kleinschmidt, et al., 2009)

Background

Bioko, Equatorial Guinea, is an island in the Gulf of Guinea with approximately 200,000 inhabitants. The Bioko Island Malaria Control Project (BIMCP) began in 2003 with funding from a consortium led by Marathon Oil Corporation and from the Government of Equatorial Guinea. It was intended to be a comprehensive, full-coverage, malaria control program. BIMCP started implementation of malaria control activities in March 2004 with a comprehensive indoor residual spraying (IRS) program. In March 2005, the program introduced artemisinin combination therapies (ACT), artesunate in combination with sulfadoxine-pyrimethamine (SP), free of charge for children under age 15 years and pregnant women, as well as intermittent preventative treatment (IPTp) with SP. In October 2007, the program initiated a vector control component through a universal insecticide-treated net (ITN) distribution and hanging campaign. BIMCP promoted adherence to these interventions through information, education, and communication campaigns.

Evaluation Study Design

The evaluation used a plausibility study design to assess changes in primary outcome measures for a period before and after the start of the BIMCP. The assumed counterfactual was that no change in primary evaluation outcomes would have occurred without the BIMCP implementation.

Primary Evaluation Outcomes

- Proportion of all households sprayed by IRS
- Proportion of households possessing at least one ITN
- Proportion of children ages 2 to 5 years old living in an IRS-treated house or sleeping under an ITN
- Mean prevalence of infection with *P. falciparum* in children ages 2 to 5 years
- Prevalence of moderate to severe anemia (less than 11 g/dL) in children ages 2 to 5 years
- All-cause under-5 mortality rate

Primary Data Sources

A system of sentinel areas, which included 86% of all houses in Bioko (according to a recent census carried out by BIMCP) formed the basis of surveillance. Surveys were carried out on random cross sections of households using a survey instrument adapted from the Malaria Indicator Survey developed by the Roll Back Malaria Monitoring and Evaluation Reference Group. Evaluators measured the hemoglobin (HemoCue, Ängelholm, Sweden) of children ages 2 to younger than 15 years and tested them for *P. falciparum* using malaria rapid tests (R&R, Cape Town, South Africa).

Window traps monitored mosquito vector species, relative abundance, sporozoite prevalence, and molecular markers of insecticide resistance. The weather station at Malabo Airport, Bioko, provided total monthly rainfall.

Data Synthesis and Analysis

The evaluators used survival regression to measure difference in all-cause, under-5 mortality. Analysts included a covariate (year) to test the presence of a declining trend in child mortality. Household access to electricity was an indicator of household wealth. Evaluators tested the association between rainfall in the previous 12 months and child mortality using a subsample. Variables of interest included possession and use of ITNs, IRS coverage, and pregnant women protected via IPTp. Analysts derived annual ACT consumption from supply chain indicators. Analysts also calculated odds ratios yearly for fever, parasitemia, and anemia rates in children ages 2–5 years.

Primary Findings

Four years after multiple high-coverage, malaria control interventions were introduced on Bioko Island, Equatorial Guinea, evaluators estimated and assessed changes in infection with malarial parasites, anemia, and fever history in children relative to changes in all-cause under-5 mortality. There were reductions in prevalence of infection (odds ratio [OR] = 0.31, 95% confidence interval [CI] = 0.2–0.46), anemia (OR = 0.11, 95% CI = 0.07– 0.18), and reported fevers (OR = 0.41, 95% CI = 0.22–0.76) in children. Under-5 mortality fell from 152 per 1,000 births (95% CI = 122–186) to 55 per 1,000 (95% CI = 38–77; hazard ratio = 0.34 [95% CI = 0.23–0.49]).

Primary Conclusions and Issues for Interpretation

The evaluation adequately accounted for non-intervention factors associated with child mortality, parasite rates, and anemia rates. During the intervention period, child mortality was greatly reduced; it is plausible that malaria control contributed too much of the observed reduction.

Appendix G. Case Study 2: Zambia (Chizema-Kawesha, et al., 2010)

Background

Zambia has an estimated population of 12 million people. It is a fully endemic country with regular and moderate to high transmission of malaria in every district; high transmission between December and May is associated with the rainy season. Zambia's National Malaria Control Strategic Plan initially covered 2000 to 2005, and was updated for 2006 to 2011. The Zambia Ministry of Health National Malaria Control Center (NMCC) and partners implemented a comprehensive, national full-coverage malaria control program consisting of the following interventions: insecticide-treated mosquito nets (ITNs), indoor residual spraying (IRS), intermittent preventive treatment in pregnancy (IPTp), rapid diagnostic tests (RDTs) and case management with artemisinin combination therapies (ACTs). This evaluation covers the scale-up efforts in Zambia from 2000 through 2008.

Evaluation Study Design

This evaluation used a plausibility study design to assess changes in primary outcome measures for a period before and after the start of the Zambia National Malaria Control Program. The assumed counterfactual was that no change in primary evaluation outcomes would have occurred without the Zambia National Malaria Control Program implementation.

Primary Evaluation Outcomes

- Proportion of all households sprayed by IRS
- Proportion of households possessing at least one ITN
- Proportion of children under 5 years of age living in a house with IRS or at least one ITN
- Mean prevalence of infection with *P. falciparum* in children ages 1–59 months
- Prevalence of severe anemia (less than 8 g/dL) in children ages 6–59 months
- All-cause post-infant mortality rate
- All-cause under-5 mortality rate

Primary Data Sources

The Zambia National Malaria Control Center (NMCC) monitored commodity procurement and distribution. The NMCC also tracked progress using programmatic output information on service delivery for malaria interventions and use of RDTs. Evaluators considered the following household surveys: the national UNICEF Multiple Indicator Cluster Survey (1999), Demographic and Health Surveys (2001-02 and 2007), Malaria Indicator Surveys (2006 and 2008), and a Roll Back Malaria survey (2004).

Data Synthesis and Analysis

Analysts measured percentage changes in parasite and anemia rates between surveys. They did not include other covariates, such as wealth. The analysis also did not include rainfall, secular trend, and other child health interventions. Separate analyses suggested that rainfall and coverage of other child health interventions did not change over the intervention time period.

Primary Findings

Zambia malaria financial support expanded from US$9 million in 2003 to US$ ~40 million in 2008. High malaria prevention coverage was achieved and extended to poor and rural areas. Increasing coverage was consistent in time and location with reductions in child (ages 6–59 months) parasitemia and severe anemia (53% and 68% reductions, respectively, from 2006 to 2008) and with lower post-neonatal infant and 1–4 years of age child mortality (38% and 36% reductions between 2001–2 and 2007 survey estimates).

Primary Conclusions and Issues for Interpretation

The rapid national scale-up of malaria control interventions outlined in the National Malaria Control Strategic Plan was associated with declining malaria parasite and anemia prevalence. The results are limited. Adjusting primary outcome changes in statistical models with the inclusion of contextual factors known to be associated with malaria parasitemia and anemia, such as wealth and rainfall, would improve the evaluation. Also, the analysis did not account for existing trends in child mortality, which was known to be declining just before the intervention period.

Appendix H. Case Study 3: Zanzibar, North A District (Bhattarai, et al., 2007)

Background

The study took place in North A District, Zanzibar, situated just off the coast of mainland Tanzania. The district is rural and has a population of about 85,000. Subsistence farming and fishing are the main occupations. *Plasmodium falciparum* is the predominant malaria species and *Anopheles gambiae* complex is considered the main vector. Malaria transmission is stable with seasonal peaks related to rainfall in March–May and October–December. Malaria transmission in the district before the interventions is reportedly high, but entomological data are not available to allow a precise characterization of malaria transmission intensity. However, *P. falciparum* prevalence of about 30% was observed in febrile children under 5 years old, suggesting that North A District had been a high transmission area before artemisinin combination therapy (ACT) implementation in September 2003. Zanzibar, including North A District, implemented ACT for uncomplicated malaria in late 2003 and insecticide-treated nets (ITNs) from early 2006. ACT was available free of charge to all malaria patients, while ITNs were distributed free to children under 5 years old and pregnant women. The study reports trends in *P. falciparum* prevalence and malaria-related health parameters following the implementation of ACTs and ITNs in North A District.

Evaluation Study Design

This evaluation used a pre- and post-intervention study design to assess changes in primary outcome measures for a period before and after the start of the scale-up of malaria control interventions.

Primary Evaluation Outcomes

- Proportion of children under 5 years of age sleeping under an ITN
- *P. falciparum* parasitemia (and gametocytemia) in children 0–14 years of age
- Outpatient malaria diagnosis, malaria-related hospital admissions, and malaria-attributed mortality
- Crude under-5 mortality

Primary Data Sources

Three cross-sectional surveys conducted in North A District between 2003 and 2006 provide estimates of *P. falciparum* parasitemia and ITN use. Health Facility Records from 13 facilities in North A District provide data on outpatient malaria diagnosis, malaria-related hospital admissions, and malaria-attributed mortality for the period 2000–2004. Records of vital events, for the period 1998–2005, obtained from the North A District Commissioner's Office, provide estimates of crude

under-5 mortality. Official registers of the Tanzania Metrological Agency of the Ministry of Communications and Transport provided total monthly rainfall.

Data Synthesis and Analysis

Analysts used logistic regression model with robust standard error to adjust for the effect of age, sex, sleeping under an ITN, asset index and survey year on asexual *P. falciparum* parasitemia and gametocytemia in the survey years. They computed relative changes (as ratios) in outpatient malaria diagnosis, malaria-related hospital admissions, malaria-attributed mortality and crude under-5 mortality. Analysts calculated correlation coefficients to assess the linear relationships between monthly rainfall and outpatient malaria diagnosis, and malaria-attributed deaths.

Primary Findings

P. falciparum parasitemia decreased in children under 5 years old between 2003 and 2006; using 2003 as the reference year, odds ratios and 95% confidence intervals were, for 2005, 0.55 (0.28–1.08), and for 2006, 0.03 (0.00–0.27); $p<0.001$for trend. Between 2002 and 2005, crude under-5, infant, and child mortality decreased by 52%, 33%, and 71%, respectively. Malaria-related admissions, blood transfusions, and malaria-attributed mortality decreased significantly by 77%, 67%, and 75%, respectively, between 2002 and 2005 in children under 5 years old. Positive correlations between monthly rainfall, outpatient malaria diagnoses and malaria-attributed deaths observed during the pre-intervention period were not observed during the post-intervention period.

Primary Conclusions and Issues for Interpretation

Following deployment of ACTs in North A District 2003, malaria-associated morbidity and mortality decreased within two years. Further distribution of ITNs in early 2006 resulted in a 10-fold reduction of *P. falciparum* parasitemia. The evaluation findings show only short-term trends in malaria control in North A District. And addition of other contextual factors associated with child survival would have improved the evaluation.

Appendix I. Case Study 4: Mainland Tanzania (Tanzania Malaria Impact Evaluation Research Group, available at www.pmi.gov)

Background

Tanzania's Ministry of Health and National Malaria Control Program (NMCP) and member organizations of the Roll Back Malaria partnership conducted an evaluation to assess the impact of the scale-up of malaria control interventions on malaria morbidity and all-cause under-5 mortality between 1999 and 2010 in Mainland Tanzania. Tanzania had a population of 41.9 million in the Mainland and approximately 1 million in the islands of Zanzibar in 2010. Mainland Tanzania and Zanzibar have different malaria transmission patterns, and each have their own malaria control program; therefore, the evaluation focused on Mainland Tanzania alone (a separate evaluation for Zanzibar is underway).

Ninety-three percent of the population of Mainland Tanzania lives in areas at risk for malaria, where *Plasmodium falciparum* is responsible for 96% of malaria infections and *Anopheles gambiae* and *An. funestus* are the primary vectors. The NMCP coordinates its malaria control activities through the Malaria Medium-Term Strategic Plans (2002–7 and 2008–13). The NMCP scaled up insecticide-treated nets (ITNs), intermittent preventive treatment in pregnant women (IPTp), indoor residual spraying (IRS) in target areas and malaria case management. Rapid tests were not scaled up until after the evaluation period of 1999–2010; therefore, the evaluation primarily based malaria diagnosis on clinical diagnosis, with microscopy restricted to hospitals during this period. Mainland Tanzania replaced the failing antimalarial drug chloroquine with sulphadoxine-pyrimethamine (SP) in 2001, and then replaced this with ACTs (artemether lumefantrine) in 2006 (by 2003 the SP failure rate was 15% at day 14). This evaluation assessed trends in malaria control intervention coverage, malaria morbidity (anemia and malaria parasite prevalence), and all cause child mortality (ACCM) during the period between 1999 and 2010. In addition, the evaluation examined changes in other socioeconomic and child health interventions during this evaluation period, which may have contributed to the declines in child mortality.

Evaluation Study Design

The evaluation used a pre- and post-intervention plausibility evaluation design that measured changes in malaria control intervention coverage, malaria-related morbidity, and ACCM, and accounted for other contextual determinants of child survival.

Primary Evaluation Outcomes

- Proportion of all households with at least one ITN
- Proportion of children under 5 years of age sleeping under an ITN

- Proportion of children under 5 years of age with fever in the past 2 weeks who received antimalarial treatment according to national policy within 24 hours from onset of fever
- Proportion of pregnant women who received intermittent preventive treatment for malaria during antenatal care clinic visits during their last pregnancy
- Proportion of children ages 6–59 months with hemoglobin less than 8g/dL
- Proportion of children ages 6–59 months with a positive rapid diagnostic test (RDT) or microscopy result
- All-cause under-5 mortality rate

Primary Data Sources

The evaluation used a series of nationwide household surveys as primary data sources: 1999 DHS, 2004–5 DHS, 2007–8 MIS and 2010 DHS. Other data sources included: the Tanzania National Household Budget Survey (2000–1 and 2007); economic reports from the Bank of Tanzania; Ifakara DSS data; NMCP parasitemia surveys in 2006 and 2008; data from the IPTp trial in Lindi and Mtwara Regions; and facility data from the Ifakara Designated, Bagamoyo, Nyakahanga, Chato and Rubya District Hospitals. Official registers of the Tanzanian Meteorological Agency provided total monthly rainfall data, and the International Research Institute for Climate and Society provided other rainfall data.

Data Synthesis and Analysis

Changes in malaria intervention coverage, malaria morbidity, all-cause mortality and contextual factors at the national level were assessed over the evaluation period (1999–2010). The analysis of anemia and mortality was stratified by age (children 6–23 months of age [those traditionally most likely to develop symptoms, illness and death from malaria in high malaria burden settings] and 24–59 months of age for comparison), by residence (urban/rural), and by malaria risk tercile. Regions were split into malaria risk terciles (higher, medium, and lower) based on the regional malaria *P. falciparum* prevalence in the 2007–8 MIS.

Primary Findings

Funding for malaria control in Mainland Tanzania over the period 2000–2010 was $450 million. Household ITN ownership increased from 22.5% (95% CI: 20.5%–24.6%) in 2004–5 to 63.5% (CI: 61.7%–65.2%) in 2010. ITN use by children under 5 years of age increased from 1.8% (CI: 1.2%–2.8%) in 1999 to 63.9% (CI: 61.2%–66.5%) in 2010. Coverage of IPTp increased slightly from 20.8% (CI: 19.0%–22.7%) in 2004–05 to 25.7% (CI: 23.6%–28.0%) in 2010. Anemia fell from 11.1% (CI: 10.0%–12.3%) in children 6–59 months of age in 2004–5 to 5.5% (CI: 4.7%–6.4%) in 2010. The relative decline in anemia in children 6–23 months of age was greater in higher-risk areas (60%) compared with medium- and lower-risk areas (28% and 17%, respectively). A 45% relative

reduction in ACCM from 148 to 81 deaths per 1,000 live births was observed, according to the 1999 and 2010 surveys. The relative decline was greater in children 6–23 months of age (49%) compared with children 24–59 months of age (34%), and greatest in children 6–23 months of age from medium-risk areas (54%) compared with higher- and lower-risk areas (41% and 37%, respectively). Several contextual factors significantly increased between 1999 and 2010, including GDP per capita, vitamin A supplementation, and early and exclusive breast feeding. At the same time, several contextual factors worsened, including pregnant women's attendance at four or more antenatal care visits and coverage with two or more doses of the tetanus toxoid vaccine.

Primary Conclusions and Issues for Interpretation

ACCM is, by definition, multifactorial, produced directly and indirectly by causes that interact with each other in myriad complex ways. A plausibility argument allows evaluators to organize a multiplicity of evidence in a causal pathway that helps them visualize and understand the potential direction and magnitude of the change studied. It also permits evaluators to draw conclusions about the impact that, in the absence of absolute proof of causality, are reasonable and based on verifiable evidence. The evidence shows that ACCM declined in Tanzania during the evaluation period 1999–2000. During that same period, socioeconomic status improved, vitamin A supplementation grew exponentially, and early and exclusive breastfeeding improved. The potential impact of all of these changes on ACCM does not account for the entirety of the decline; therefore, it is reasonable to conclude that the scale-up of malaria control interventions influenced at least part of the remaining impact.

Limitations of the plausibility argument include the lack of a counterfactual, no statistical proof that malaria interventions and ACCM decline are linked, and the possibility that not all contextual or confounding factors have been accounted for. Counterfactuals in national-level impact evaluations are almost impossible to find. In the absence of a counterfactual, statistical proof of impact is difficult to generate. However, statistical models that use operational and other program data showing known links between interventions and outcome variables could have strengthened the plausibility argument in the Tanzania evaluation. The amount and types of data, including that there was only one nationally representative parasitemia estimate available and only some health facility data available during the evaluation period, limit this evaluation in Tanzania. In addition, the timing of the mortality and intervention coverage estimates are not completely aligned. The follow-up mortality estimate from the 2010 DHS survey refers to the 2006–2010 period (2008 mid-point), which is more in line with the 2007–8 intervention coverage estimates than the 2010 intervention coverage estimates.

Appendix J. Case Study 5: Malawi (Malawi Malaria Impact Evaluation Research Group, will be available at www.pmi.gov)

Background

In Malawi, the National Malaria Control Program (NMCP) of the Ministry of Health (MoH) and member organizations of the RBM Partnership conducted an evaluation to assess the impact of the scale-up of malaria control interventions on malaria morbidity and all-cause under-5 mortality between 2000 and 2010. Malaria is highly endemic in Malawi, with 95% of the population living in areas at risk for malaria. *Plasmodium falciparum* is responsible for 98% of malaria parasite infections and *Anopheles gambiae s.s.*, *An. Arabiensis*, and *An. funestus* are the primary malaria vectors. The MoH formally established the NMCP in 1987. The NMCP has coordinated its activities through three, 5-year Malaria Strategic Plans. The NMCP has scaled up insecticide- treated nets (ITNs), indoor residual spraying (IRS) in target areas (seven high-malaria-prevalence districts in 2010), and malaria case management.

In 1993, Malawi was the first country in Sub Saharan Africa to establish an intermittent preventive treatment for pregnant women (IPTp) policy and has scaled up this intervention ever since. Rapid diagnostic tests (RDTs) were not scaled up until after the evaluation period of 2000–2010; therefore, the evaluation primarily based malaria diagnosis on clinical diagnosis with microscopy restricted to hospitals during this period. Malawi replaced the failing antimalarial drug chloroquine with sulphadoxine-pyrimethamine (SP) in 1993 and then replaced SP with artemisinin combination therapies (ACTs), artemether lumefantrine, in 2007. This evaluation assessed trends in malaria control intervention coverage, malaria morbidity (anemia, malaria parasite prevalence, and malaria cases at health facilities), and all-cause child mortality (ACCM) during the period 2000–2010. The evaluation also examined changes in other socioeconomic and child health interventions during this evaluation period, which may have contributed to the declines in child mortality. In addition, the evaluation examined associations between ITNs and malaria morbidity and mortality under programmatic conditions in Malawi during this period.

Evaluation Study Design

The evaluation used a pre- and post-intervention plausibility evaluation design that measured changes in malaria control intervention coverage, malaria-related morbidity, and ACCM—and accounted for other contextual determinants of child survival. Analyses that assessed the association between ITN ownership and anemia, malaria parasite prevalence, severe malaria cases, and child survival further supported the plausibility design.

Primary Evaluation Outcomes

- Proportion of all households with at least one ITN
- Proportion of children under 5 years of age sleeping under an ITN
- Proportion of children under 5 years of age with fever during the past 2 weeks who received antimalarial treatment according to national policy within 24 hours from onset of fever
- Proportion of pregnant women who received intermittent preventive treatment for malaria during antenatal care clinic visits during their last pregnancy
- Proportion of children ages 6–59 months with hemoglobin less than 8g/dL
- Proportion of children ages 6–59 months with a positive RDT or microscopy result
- All-cause under-5 mortality rate

Primary Data Sources

The evaluation used a series of nationwide household surveys as primary data sources: 2000 DHS, 2004 DHS, 2006 MICS, and 2010 DHS. Other data sources included the 2010 MIS, national micronutrient surveys (NMS) from 2001 and 2009, a series of subnational anemia and parasitemia surveys from 2005–2009, commodity data from the Logistics Management Information System (LMIS), and programmatic data on ITN distribution. The evaluation also looked at health facility data from HMIS and IDSR from 2005–2010. NASA Land Processes Distributed Active Archive Center Data Pool from the USGS/Earth Resources Observation and Science Center, MODIS satellite (temperature), and USGS FEWS NET data portal (rainfall) provided weather data.

Data Synthesis and Analysis

The evaluation assessed changes in malaria intervention coverage, malaria morbidity, all-cause mortality, and contextual factors at the national level over the period from 2000 to 2010. Evaluators stratified the analysis of anemia and mortality by age (children 6–23 months of age [those traditionally most likely to develop symptoms, illness and death from malaria in high malaria burden settings] and children 24–59 months of age for comparison), by residence (urban or rural) and by malaria-risk tercile. Clusters from the 2010 DHS survey were split into malaria-risk terciles (higher, medium, and lower) based on the *P. falciparum* prevalence in children 2–10 years old ($PfPR_{2-10}$) from the 2007 Malaria Atlas Project map. The analysis used several statistical models to strengthen the plausibility argument as summarized in Table J.1.

TABLE J.1: MULTIVARIABLE REGRESSION MODELS USED IN THE MALAWI IMPACT EVALUATION

Type of Analysis	Hypothesis Tested	Data Sources
Random-effects logistic regression model	Has increasing ITN ownership led to reductions in anemia?	A&P surveys [subnational, 6-30 mos.] (ITNs & anemia); MODIS satellite (temperature); FEWS NET (rainfall)
Random-effects logistic regression model	Has increasing ITN ownership led to reductions in malaria parasitemia?	A&P surveys [subnational, 6-30 mos.] (ITNs & parasitemia); MODIS satellite (temperature); FEWS NET (rainfall)
Random-effects Poisson regression model	Has increasing ITN ownership led to reductions in severe malaria cases?	Program distribution data (ITNs); IDSR (severe inpatient malaria cases); 1998 and 2008 census (mid-year, district-level population data); MODIS satellite (temperature); FEWS NET (rainfall)
Cox proportional hazards model (matched strata included as a shared frailty)	Is ITN ownership protective against mortality in children under five years of age?	2010 DHS (ITNs, mortality, other covariates); MAP 2010 ($PfPR_{2-10}$); MODIS satellite (temperature); FEWS NET (rainfall)
District Level Poisson regression	Has the scale-up of ITN ownership led to declines in mortality in children under five years of age?	2006 MICS & 2010 DHS (ITNs, mortality, other covariates); MAP 2010 ($PfPR_{2-10}$); FEWS NET (rainfall)

Primary Findings

Household ITN ownership increased from 27.4% (95% CI: 25.9%–29.0%) in 2004 to 56.8% (CI: 55.6%–58.1%) in 2010. ITN use by children under 5 years of age increased from 2.8% (CI: 2.2%–3.4%) in 2000 to 39.4% (CI: 38.0%–40.8%) in 2010. Coverage of IPTp increased from 42.9% (CI: 41.1%–44.8%) in 2004 to 53.8% (CI: 52.2%–55.5%) in 2010. Nationally, anemia did not significantly change between 2004 (10.6%, CI: 9.1%–12.4%) and 2010 (8.7%, CI: 7.6%–10.0%); however, anemia declined 36% in children 6–23 months of age, but did not change significantly in children 24–59 months of age, between 2004 and 2010. There was a significant decline in anemia in children 6–23 months of age in higher and medium malaria risk areas but not in lower-risk areas. During the dry season, malaria parasitemia decreased from 60.5% (CI: 53.0%–68.0%) in 2001 to 20.4% (CI: 15.7%–25.1%) in 2009 in children 6–35 months of age.

The number of suspected malaria cases in children under 5 years and all ages increased between 2005 and 2010; however, this finding was not adjusted for increased reporting and increased care seeking. A 41% relative reduction in ACCM from 188 to 112 deaths per 1,000 live births was observed, according to the 2000 and 2010 surveys. The logistic regression models showed that household ITN ownership was protective against malaria parasitemia (OR=0.81, CI: 0.72–0.92) and anemia (OR=0.77, CI: 0.70–0.86) after controlling for other covariates. ITN distributions per 1,000 population were not significantly associated with inpatient malaria cases at facilities included in IDSR, according to the random-effects Poisson regression model. The Cox Proportional Hazards Model revealed that household ITN ownership was associated with a significant reduction in

mortality risk for children 1–59 months of age (HR=0.75, CI: 0.62–0.90) after adjusting for other covariates. District-level household ITN ownership was significantly associated with child survival (IRR=0.55, CI: 0.31–0.99) after controlling for other covariates based on the district-level Poisson regression model.

Primary Conclusions and Issues for Interpretation

The impact evaluation in Malawi used a plausibility evaluation design to conclude that the scale-up of malaria control interventions in Malawi contributed to the decline in ACCM observed between 2000 and 2010. In addition to analyzing and discussing the potential effects of socioeconomic and child health interventions (contextual factors) on the declines in ACCM and an assessment of health facility data, this evaluation added several types of statistical modeling to strengthen the plausibility evaluation design and account for potential confounding. The models revealed that ITN ownership was protective against anemia and parasitemia and was significantly associated with reductions in mortality risk for children 1–59 months of age under operational conditions in Malawi. The Malawi evaluation is an example of where models were added to further support or disprove the plausibility argument. A limitation of this evaluation is that the timing of the mortality and intervention coverage estimates are not completely aligned. The follow-up mortality estimate from the 2010 DHS survey refers to the 2006–2010 period (2008 mid-point), which precedes the 2010 intervention coverage estimates. In addition, mortality declines were already apparent before the scale-up of malaria control interventions.