Framework for Evaluating National Malaria Programs in Moderate- and Low-Transmission Settings: *Aide Memoire*

April 2019

USAID FROM THE AMERICAN PEOPLE
CDC
U.S. President's Malaria Initiative
RBM Partnership To End Malaria
MEASURE Evaluation

Framework for Evaluating National Malaria Programs in Moderate- and Low-Transmission Settings: *Aide Memoire*

April 2019

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1-919-445-9359
measure@unc.edu
www.measureevaluation.org

This publication has been supported by the President's Malaria Initiative (PMI) through the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AIDOAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of PMI, USAID, or the United States government. TR-19-343
ISBN: 978-1-64232-140-1

USAID FROM THE AMERICAN PEOPLE
CDC
U.S. President's Malaria Initiative
RBM Partnership To End Malaria
MEASURE Evaluation

Acknowledgments

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

Yazoumé Yé (MEASURE Evaluation, ICF), Samantha Herrera (MEASURE Evaluation, ICF), Andrew Andrada (MEASURE Evaluation, ICF), Debra Prosnitz (MEASURE Evaluation, ICF), Ruth Ashton (MEASURE Evaluation, Tulane University School of Public Health and Tropical Medicine), Thom Eisele (MEASURE Evaluation, Tulane University School of Public Health and Tropical Medicine), Achuyt Bhattarai (President's Malaria Initiative [PMI]/United States Centers for Disease Control and Prevention [CDC]), Erin Eckert (PMI/United States Agency for International Development [USAID]), Lia Florey (PMI/USAID), Arantxa Roca-Feltrer (Malaria Consortium), Chris Drakeley (London School of Hygiene and Tropical Medicine), Sean Hewitt (VBDC Consulting Ltd), Ryuichi Komatsu (the Global Fund to Fight AIDS, Tuberculosis and Malaria [Global Fund]), Estifanos Shargie (Global Fund), Roger Tine (Université Cheikh Anta Diop de Dakar), Adam Bennett (University of California San Francisco), Cameron Taylor (Demographic and Health Surveys Program/ICF), John Painter (PMI/CDC), Anna Bowen (PMI/CDC), Abdul-Wahid Al-Mafazy (Zanzibar Malaria Elimination Programme), Rebecca Kiptui (National Malaria Control Program, Kenya), Mateusz Plucinski (CDC), Peter McElroy (PMI/CDC), Christelle Gogue (PATH), Misun Choi (PMI/USAID), Frank Chacky (National Malaria Control Program, Tanzania), Manuel Hetzel (Swiss Tropical and Public Health Institute), Sumaiyya Thawer (Swiss Tropical and Public Health Institute/National Malaria Control Program, Tanzania), Tabitha Kibuka (PMI Impact Malaria Project/Population Services International), Bolanle Olapeju (VectorWorks/Johns Hopkins University), Deepa Pindolia (Clinton Health Access Initiative), Inessa Ba (Clinton Health Access Initiative), and Arnaud Le Menach (Clinton Health Access Initiative). The Evaluation Task Force thanks all other individuals who contributed at various stages in the development of this framework.

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: Agbessi Amouzou and Melissa Marx (both of the Johns Hopkins Bloomberg School of Public Health) and Alexander Rowe (CDC).

Special thanks also go to Cindy Young-Turner for editing and proofreading multiple versions of this document.

We thank the knowledge management team of MEASURE Evaluation, University of North Carolina at Chapel Hill, for editorial, design, and production services.

Suggested citation

Evaluation Task Force of Roll Back Malaria's Monitoring and Evaluation Reference Group. (2019). *Framework for Evaluating National Malaria Programs in Moderate and Low Transmission Settings: Aide Memoire*. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina.

Contents

Acknowledgments ii
Abbreviations iv
1. Background 5
2. Measuring national malaria program achievements 6
3. Evaluation design 9
4. Gathering evidence 10
5. Data analysis, synthesis, and interpretation 11
6. Implementing the evaluation framework 12
7. Conclusions 13
8. References 14

Abbreviations

CHIS	community health information system
CRVS	civil registration and vital statistics system
HDSS	health and demographic surveillance system
ITS	interrupted time series
NMP	national malaria program
NMSP	national malaria strategic plan
RHIS	routine health information system

1. Background

The epidemiology of malaria has become increasingly heterogeneous in many countries. These countries require granular data on transmission risk and incidence to effectively inform and target their interventions and track their progress. To meet these needs for strengthening national malaria programs (NMPs), an evaluation task force comprising a subgroup of the Roll Back Malaria Monitoring and Evaluation Reference Group developed the *Framework for Evaluating National Malaria Programs in Moderate- and Low-Transmission Settings*. This document provides an overarching framework for evaluating NMPs along the continuum of malaria transmission. Each transmission setting is defined by the World Health Organization classifications [1]. The scope and objectives were informed through a review and synthesis of existing guidance documents and tools for malaria surveillance, monitoring, and evaluation. This aide memoire summarizes the larger framework document.

The key objectives of the evaluation framework are to provide the following:

- An overarching framework for evaluating NMPs along the continuum of malaria transmission
- Description of linkages between impact and process evaluation
- Specific recommendations and guidance for conducting impact evaluations in countries with moderate-, low-, and heterogeneous-transmission settings
- Guidance on how to bring together evaluation results at the subnational level to tell a national-level narrative in heterogeneous-transmission settings

3. Measuring national malaria program achievements

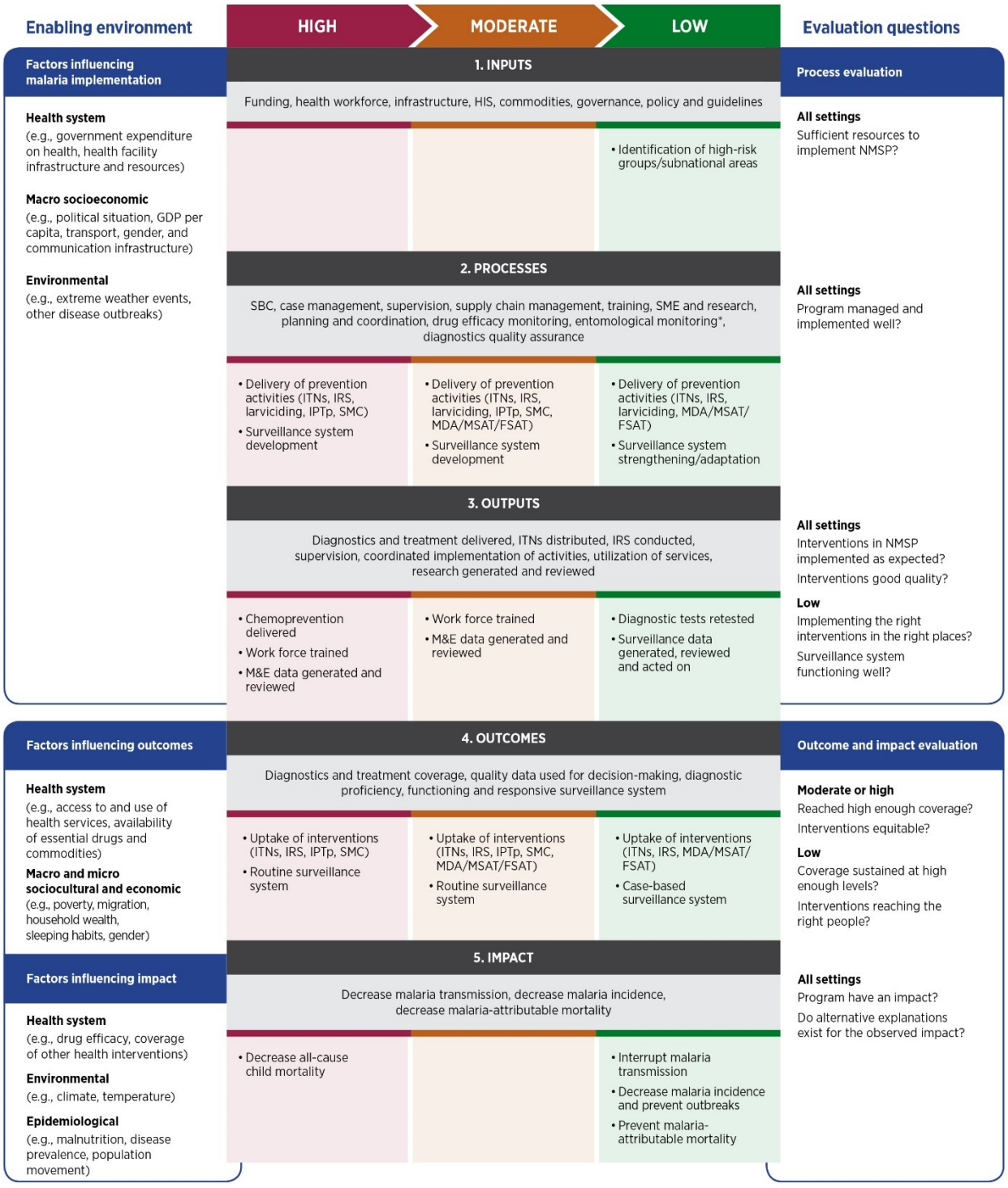
The framework theory of change for NMPs (Figure 1) summarizes key evaluation objectives, questions, and indicators across transmission settings. In addition, it outlines the relationships between the inputs, processes, and outputs of a malaria program, with the expected outcomes and desired impact of the program. In all transmission settings, NMPs aim to reduce the number of malaria cases and deaths. In high-transmission settings, it is not feasible to count malaria deaths; therefore, all-cause child mortality is used as a proxy measure of impact on malaria mortality. In low- and moderate-transmission settings, where fewer deaths are attributable to malaria, incidence is the primary indicator of impact.

In all transmission settings, attaining high coverage of key interventions and prompt and effective case management is critical to achieving impact. In addition, malaria surveillance data of adequate quality are important in the selection of evidence-based intervention packages. In low-transmission settings, a strong, functioning, and responsive surveillance system will become increasingly critical to inform evidence-based decision making, including effective targeting of interventions, to achieve the program's desired impact. This is an iterative process informed by the evaluation of NMPs.

Several contextual factors may affect program implementation, outcomes, and impact. Contextual factors may be sociocultural, economic, environmental, epidemiological, or related to the health system. The theory of change highlights at which point in the malaria program cycle certain types of contextual factors may be relevant to consider.

Evaluations of NMPs provide critical information for programmatic and policy decision making. A process evaluation assesses the degree to which an NMP (and its national malaria strategic plan [NMSP]) has been implemented—and why or why not. A process evaluation examines questions, such as whether sufficient inputs have been allocated or mobilized for a program, what activities have been undertaken, and who has been reached by the program activities. In high- and moderate-transmission settings, process evaluations focus on assessing the full package of interventions to identify bottlenecks and improve program implementation. In low-transmission settings that are marginally conducive for malaria transmission, the process evaluation will emphasize assessment of case management implementation and performance of the surveillance system. Process evaluations can characterize the strength of program implementation and link program inputs to achieved outcomes.

Figure 1. Theory of change for national malaria programs across the transmission spectrum



*May include net durability monitoring and IRS application quality monitoring using cone bioassay. WHO provides specific guidance on entomological surveillance, and intervention monitoring and evaluation available at **http://www.who.int/malaria/areas/vector_control/entomological_surveillance/en/**

Definitions: **NMSP**=national malaria strategic plan; **ITNs**=insecticide-treated nets; **IRS**=indoor residual spraying; **IPTp**=intermittent preventive treatment in pregnancy; **SMC**=seasonal malaria chemoprevention; **SME**=surveillance, monitoring, and evaluation; **SBC**=social behavior change; **HIS**=health information system; **MDA/MSAT/FSAT**=mass drug administration/mass screening and treatment/focal screening and treatment

An outcome evaluation determines whether the program reached the expected intervention coverage at the population level. In high- and moderate-transmission settings, outcome evaluations focus on assessing the coverage level of interventions achieved across subpopulations (e.g., children under five, rural/urban). In low-transmission settings, the focus is on assessing the coverage of diagnostics and treatment for malaria, in addition to assessing the intervention coverage among the populations at risk for malaria. It is important to conduct process and outcome evaluations to better inform impact evaluations.

An impact evaluation assesses whether the program had an effect on malaria transmission and malaria-attributable morbidity and mortality, and whether it achieved its goals. Specifically, an impact evaluation aims to assess the changes in impact measures that can be attributed to a particular package of interventions implemented by the NMP. Impact evaluations must be tailored to the country context, and contextual factors must be accounted for because they can confound the association between the program and its potential impact. In this document we refer to impact evaluations as evaluations that assess the plausible contribution of program interventions to malaria health outcomes and the changes in malaria health outcomes.

In high-transmission settings, all-cause child mortality is the primary recommended impact indicator. In moderate- and low-transmission settings, malaria case incidence is the primary recommended indicator. Secondary indicators for moderate- and low-transmission settings include malaria test positivity rate, proportion of malaria admissions, malaria mortality, number of annual malaria outbreaks, parasite prevalence, and seroprevalence. The larger framework document offers a complete list of monitoring and evaluation indicators for malaria across high-, moderate-, and low-transmission settings, along with an indicator reference guide in Annex 2. Impact evaluations should be prospectively planned and include process and outcome evaluations. These evaluations should be conducted toward the end of the NMSP cycle and include data from the previous NMSP implementation to fully understand the plausible attribution of impact to the NMP.

5. Evaluation design

The evaluation design is influenced by the priority questions of the NMP and other stakeholders, transmission settings of the country, data sources available, data quality, interventions applied, and strategies used to introduce or scale up these interventions. Evaluation designs most suited to moderate- and low-transmission settings are quasi-experimental designs, which use non-randomized exogenous variation in the exposure of interest to estimate effect sizes. Compared to observational studies, the rigorous design and analytical methods allow quasi-experimental studies to better account for threats to internal validity. Interrupted time series (ITS) analyses, a type of quasi-experimental design, have been shown to be particularly strong [2, 3]. The use of counterfactuals, which describe the outcome in the absence of the program being evaluated, are important for establishing the impact of the program. Experimental methods, such as randomized controlled trials, estimate the counterfactual from the control group. Quasi-experimental and observational studies use various methods and assumptions to estimate the counterfactual.

The transmission strata of a country should be considered in the impact evaluation design. This may involve conducting the analysis within each stratum to understand the stratum-specific program impact. Alternatively, it may be appropriate to use different impact evaluation analyses within each stratum because the interventions applied and data available may differ between strata. Local-level stratification from intervention deployment should not be considered in the impact evaluation design. Other considerations include high- and low-risk populations, statistical power, and scale of NMP activities and programs. Countries without an existing stratification (but known to have heterogeneous transmission) may use an interim approach, which uses baseline impact indicator values (e.g., confirmed malaria incidence) to define levels for a subgroup or stratified analysis.

Data availability and quality are important considerations for determining the evaluation design and interpreting results. In settings where an evaluation is being prospectively planned, the preferred evaluation design can inform the types of data collected. Some analytical approaches require data to be collected over the evaluation period (e.g., ITS), and others can use cross-sectional survey data (e.g., difference-in-differences). In addition, indicator definitions or reporting methods may change over time, and any such changes should be considered in the evaluation design. This is particularly relevant if routine surveillance data are used. Data quality is important to consider because use of poor-quality data can result in misleading or incorrect evaluation findings. Data do not need to be perfect; "adequate quality" data may be sufficient, and some issues with data quality can be accounted for in the analysis. Hypothetical scenarios addressing some of these common challenges are presented in the full framework document.

The interventions and NMP strategy for scaling up these interventions should also be considered in the evaluation design, because each has its own challenges and potential design solutions. Examples include a phased approach to scale up an intervention, policy changes that rapidly affect which interventions are rolled out, or situations in which the intervention or program was implemented everywhere and there are no clearly defined "control" areas. In designing the impact evaluation analysis approach and identifying data sources, including multiple data sources, running multiple analyses, and using a range of techniques to address gaps and biases in the data can improve the plausibility of findings from quasi-experimental evaluation approaches [4]. This is often described as triangulation.

7. Gathering evidence

Gathering evidence for evaluation entails defining evaluation indicators, identifying relevant data sources, gathering data sets, assessing data quality, and analyzing, triangulating, and interpreting the data. Key data sources used for intervention coverage, impact measures, and contextual factors are routine health information systems (RHIS), community health information systems (CHIS), surveys, health and demographic surveillance systems (HDSS) or sentinel sites, verbal autopsy, civil registration and vital statistics systems (CRVS), and entomological surveillance. Each data source has its own strengths and limitations that should be considered when the data gathered are interpreted. Taking into account these key data sources, strengths, limitations, and relevant stratifications, the recommended impact indicators for moderate- to low-transmission settings are shown in Table 1.

Table 1. Recommended impact indicators for moderate- and low-transmission settings

Indicator	Transmission setting		Data sources
	Moderate	Low	
Malaria case incidence: number and rate per 1,000 people per year*	X	X	RHIS, CHIS, HDSS/sentinel sites
Malaria test positivity rate*	X	X	RHIS, CHIS, HDSS/sentinel sites
Proportion of admissions for malaria	X	X	RHIS, HDSS/sentinel sites
Malaria mortality: number and rate per 100,000 people per year	X	X	RHIS, HDSS/sentinel sites, CRVS
Proportion of inpatient deaths due to malaria	X	X	RHIS, HDSS/sentinel sites
All-cause child mortality (number of child deaths per 1,000 live births)	X		Population-based household survey (DHS, MICS), census data, CRVS, HDSS/sentinel sites
Annual number of malaria epidemics	X	X	RHIS, program data
Parasite prevalence*	X	X	Population-based household survey (DHS, MIS, MICS)
Seroprevalence		X	Population-based household survey

*Disaggregated by vector species, if possible

DHS=Demographic and Health Survey, MICS=Multiple Indicator Cluster Survey, MIS=Malaria Indicator Survey

Contextual factors that may confound the association between interventions and impact indicators should be addressed in the evaluation design and analysis to assess the extent to which these factors may affect coverage and impact, and to increase the validity of the evaluation [5].

9. Data analysis, synthesis, and interpretation

The choice of analytic method should be determined by a combination of the evaluation questions, the intervention implementation approach used in the area under evaluation (e.g., phased introduction), data availability, and data quality. Evaluation methodologies relevant to impact evaluation in moderate- and low-transmission settings are ITS, dose-response, constructed controls (matching or discontinuity designs, instrumental variables), and stepped-wedge. These are not necessarily specific to low- and moderate-transmission settings only, but they demonstrate the breadth of analyses that is possible with the types of data likely to be available in these settings. Analytic techniques relevant to impact evaluation in low-transmission settings include difference-in-differences, instrumental variables, and matching to construct controls (exact, group, and propensity score matching). In addition to quantitative data, impact evaluations also require qualitative data, such as a narrative description of the program over the evaluation period, timeline of key activities, policy changes, or other contextual factors. In selecting a design and analytic approach, it is important to consider the internal and external validity of each method. Examples of a range of evaluation designs and analytic approaches are included in the full framework document, with further discussion of the advantages and limitations of each methodology.

Impact evaluations focus on examining the relationship between the NMP outcomes (e.g., coverage of interventions) and the desired impact of the program. Without process evaluation findings, why the program has or has not achieved its impact can be unclear. Process evaluations provide this critical information on why and how a program worked, and therefore are valuable in providing the necessary context to elucidate the relationships between intervention implementation and achieved outcomes and impact. In an impact evaluation of a national program, synthesizing the findings from the process evaluation and linking them to the impact evaluation becomes even more important because the evaluation examines the impact of a package of interventions, rather than the impact of just a specific intervention. Findings from a process evaluation may help determine whether limited observed impact was because of an issue in the program's theory of change or issues with program implementation.

Where impact evaluation analysis has been stratified according to differential risk areas, risk populations, or different intervention packages, compiling these individual findings into a descriptive overall narrative at the national level is often valuable. This national-level narrative may be particularly relevant to advocate continued support to and funding of the malaria program, both from national and international sources.

11. Implementing the evaluation framework

The implementation of the evaluation framework involves seven interrelated steps: (1) engaging stakeholders; (2) describing the malaria program; (3) determining the evaluation objectives, questions, and design; (4) gathering evidence; (5) analyzing the data; (6) disseminating and using the evaluation findings; and (7) improving and strengthening the NMP. Prospective planning of evaluations is encouraged to allow for greater stakeholder involvement and buy-in for evaluation, promote timely implementation, and ensure appropriate data collection. This framework is applicable for all organizations that may lead a process or impact evaluation.

A realistic timeline for carrying out the evaluation should be developed at the onset of the evaluation, to set expectations for when each stage of the evaluation will be completed and when the results will be available. On average, an impact evaluation will take 14 months to complete. To ensure that the results are relevant and useful for informing adjustments to the NMP, the evaluation should be carried out in a timely manner and ideally toward the end of the NMSP cycle.

An important part in planning for an evaluation is determining the required human resources, skills, and evaluation costs. The members of the evaluation team should be people with a solid understanding of malaria epidemiology and of the malaria program in the country, strong quantitative and qualitative research and analytic skills, knowledge of data quality dimensions and how to assess data quality, and skills in data visualization and writing evaluation findings. A detailed evaluation budget should be prepared before the evaluation is implemented. The budget should cover costs for human resources, stakeholder meetings, data collection and access to existing data as appropriate, writing, translation, printing, and dissemination.

13. Conclusions

This aide memoire summarizes the full framework document. The full document highlights the iterative process of evidence-based decision making by NMPs informed by evaluation, emphasizes the integral nature of process evaluation to impact evaluation, discusses other considerations for design such as data sources and quality, addresses ways to mitigate bias in design and analysis, and outlines timing considerations for each type of evaluation (process, outcome, and impact). This framework fills a gap for countries with low and heterogeneous malaria transmission, providing comprehensive recommendations for evaluating the impact of their NMPs as they progress.

The complete framework document is available at https://www.measureevaluation.org/resources/publications/tr-19-334/.

15. References

1. World Health Organization (WHO): **Malaria surveillance, monitoring & evaluation: A reference manual.** pp. 105. Geneva: WHO; 2018:105.

2. Cochrane Effective Practice and Organisation of Care (EPOC): **What study designs can be considered for inclusion in an EPOC review and what should they be called? EPOC resources for review authors.** Oslo: Norwegian Knowledge Centre for the Health Services; 2017.

3. Fretheim A, Zhang F, Ross-Degnan D, Oxman AD, Cheyne H, Foy R, Goodacre S, Herrin J, Kerse N, McKinlay RJ, et al: **A reanalysis of cluster randomized trials showed interrupted time-series studies were valuable in health system evaluation.** *J Clin Epidemiol* 2015, **68:**324-333.

4. Victora CG, Black RE, Boerma JT, Bryce J: **Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations.** *The Lancet* 2011, **377:**85-95.

5. Victora CG, Schellenberg JA, Huicho L, Amaral J, El Arifeen S, Pariyo G, Manzi F, Scherpbier RW, Bryce J, Habicht JP: **Context matters: Interpreting impact findings in child survival evaluations.** *Health Policy and Planning* 2005, **20 Suppl 1:**i18-i31.

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1-919-445-9359
measure@unc.edu
www.measureevaluation.org

This publication has been supported by the President's Malaria Initiative (PMI) through the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AIDOAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of PMI, USAID, or the United States government. TR-19-343
ISBN: 978-1-64232-140-1

USAID FROM THE AMERICAN PEOPLE
CDC
U.S. President's Malaria Initiative
RBM Partnership To End Malaria
MEASURE Evaluation