Research White Paper

Global Health Evidence Evaluation Framework

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • **www.ahrq.gov**

Research White Paper

Global Health Evidence Evaluation Framework

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10062-I

Prepared by:
Southern California Evidence-based Practice Center
Santa Monica, CA

Investigators:
Paul G. Shekelle, M.D., Ph.D.
Margaret A. Maglione, M.P.P.
Jill Luoto, Ph.D.
Breanne Johnsen, B.S.
Tanja R. Perry, B.H.S.

AHRQ Publication No. 13-EHC008-EF
January 2013

This report is based on research conducted by the Southern California Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10062-I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Shekelle PG, Maglione MA, Luoto J, Johnsen B, Perry TR. Global Health Evidence Evaluation Framework. Research White Paper (Prepared by the Southern California Evidence-based Practice Center under Contract No. 290-2007-10062-I). AHRQ Publication No. 13-EHC008-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2013. www.effectivehealthcare.ahrq.gov.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director, Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Christine Chang, M.D., M.P.H.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: Elizabeth Higgs, M.D., M.I.A., D.T.M.H., U.S. Agency for International Development—Washington, DC

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who participated in developing this report follows:

Robert Balster, M.D.
Jefferson Science Fellow
U.S. Agency for International Development
Washington, DC

Stef Baral, M.D., M.P.H.
Assistant Scientist
Johns Hopkins University
Baltimore, MD

Al Bartlett, M.D.
Director, Saving Newborn Lives Program
Save The Children
Washington, DC

Leonard Bickman, Ph.D.
Professor of Psychology, Psychiatry and Public Policy
Director, Center for Evaluation and Program Improvement (CEPI)
Peabody College
Vanderbilt University
Nashville, TN

Kevin DeCock, M.D.
Director, Center for Global Health
Centers for Disease Control and Prevention
Atlanta, GA

Donna Farley, Ph.D., M.P.H.
Senior Health Policy Researcher
RAND Corporation
Pittsburgh, PA

Geoff Garnett, Ph.D.
Professor of Microparasite Epidemiology
Imperial College London
London, England

Paul Gertler, Ph.D.
Professor of Economics
University of California at Berkeley
Berkeley, CA

Irene Koek, Ph.D.
Director of the Health office for U.S. Agency for International Development
Indonesia

Cliff Lane, M.D., M.P.H.
Professor of Primary Health Care
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD

Jack Molyneaux, M.D.
Director of Impact Evaluation
Millennium Challenges Corporation
Washington, DC

Pierre Ongolo-Zogo, M.D.
Head, Centre for Development of Best Practices in Health
Yaoundé Central Hospital
Yaoundé, Cameroon

Mead Over, Ph.D.
Senior Vice President
Healthcare Improvement Division
Center for Global Development
Washington, DC

Ariel Pablos-Méndez, M.D., M.P.H.
Assistant Administrator for Global Health
U.S. Agency for International Development
Washington, DC

Tikki Pangestu, Ph.D.
Director, Research Policy & Cooperation
World Health Organization
Geneva, Switzerland

Larry Slutsker, M.D., M.P.H.
Associate Director for Science
Centers for Disease Control and Prevention
Atlanta, GA

Kieran Walshe, Ph.D.
Professor of Health Policy & Management
Manchester Business School
University of Manchester
Manchester, United Kingdom

Global Health Evidence Evaluation Framework

Structured Abstract

Objectives. The Global Health Evidence Evaluation Framework project is aimed at developing an evidence framework for the purposes of informing efficacious, effective, sustainable global health programs at the community and scale level.

Data sources. Literature search of published and gray literature, input from a multidisciplinary Technical Expert Panel (TEP).

Methods. With input from the TEP we identified six existing evidence frameworks for public health/global health interventions or programs and applied these frameworks to the evidence bases for three exemplar interventions chosen to represent a diverse set of global health programs or interventions: household water chlorination, prevention of mother-to-child transmission of HIV, and lay community health workers to reduce child mortality. Review of the findings with the TEP identified an important gap as the reporting of information about the implementation of the intervention or program. We identified three existing criteria sets for implementation reporting, and selected from them 10 criteria that covered areas identified by the TEP. We pilot tested these 10 criteria on three published articles of effectiveness for each of the three exemplar interventions.

Results. Assessing the same body of evidence yields different conclusions regarding strength of evidence depending on which framework is used, some of which were extreme (with the same evidence base assessed as “strong” in one framework and “low quality” in another framework). All frameworks focus on efficacy and/or effectiveness with most attention going to the allocation method of study participants to the intervention (randomization versus other method) in assessing study quality. Many fail to consider implementation issues, and none explicitly assessed costs or sustainability at the community or scale level. Incorporating insights from the frameworks for implementation criteria helped to address some of these gaps, yet our pilot test of 10 implementation reporting criteria on nine published studies showed great variability both within and across studies. Some criteria were classified as having “good” reporting in almost all studies (criteria about the setting of the study, the characteristics of the recipients and the mode of delivery of the intervention or program) while others were reported on rarely or never (the rationale for the intervention, the costs of the intervention, an assessment of the population needs, and two criteria about barriers and facilitators of the implementation). Similarly, two articles had “good” or “fair” documentation for greater than 75 percent of criteria, while three articles had “poor or none” documentation for more than 50 percent of criteria.

Conclusions. Existing frameworks for the assessment of public health evidence do not deliver key pieces of information to inform best practices for community and large-scale global health programs, with the lack of information about implementation and sustainability being an important identified gap. In a pilot study, our application of existing criteria for the reporting of implementation information shows great variability within and across published global health intervention studies. The reporting of implementation information using criteria that have broad support across diverse stakeholders and that can be reliably assessed would provide policymakers a stronger evidence base upon which to make decisions.

Contents

Tables

Appendixes

Background

The global community has collaboratively taken responsibility for advancing health for all and codified this commitment in the Millennium Development Goals. The global effort has enabled standardization of approaches, including the agreement that evidence-based decision making is a norm for health policy, practice, and programs. The U.S. Agency for International Development (USAID) is likewise committed to evidence-based, innovative, efficacious, effective and sustainable global health programs. Evidence based decision making is critical to informing and guiding programming in global health intervention, global health policies, and programs. However, many current evidence evaluation frameworks have evolved from the clinical model where physician decision making is determined by rigorous systematic review of data derived from randomized clinical trials (RCTs) which emphasize efficacy for the individual patient.

Evidence requirements for global health programs are complex, and include three streams: efficacy at the individual level, effectiveness at the population level, and sustainability at the host country level. A standardized approach to evaluation of evidence will strengthen and support global health strategies, programs, and practice optimizing limited resources for maximal health impact.

Objective

The Global Health Evidence Evaluation Framework project is aimed at developing an evidence framework for the purposes of informing efficacious, effective, sustainable global health programs at the community and scale level.

Methods

This project represents collaboration between the project team and an interdisciplinary group of experts, called the Technical Expert Panel, or TEP. Each key step of the project requires both preparatory work on the part of the project team and then consideration and discussion by the TEP, with synthesis of the TEP discussion and decisions then made by the project team.

Preliminary Key Questions were:

- What is the role of context in the effectiveness and scalability of global health interventions/programs at the community and population level?
- What descriptions of elements of study design and execution are necessary to reach conclusions about effectiveness?
- What descriptions of elements of implementation are necessary to reach conclusions about scalability and sustainability?

The six steps of the project are to:

1. Assemble an interdisciplinary panel of experts.
2. Select a set of "exemplar interventions or programs." Note "exemplar" means "characteristic of its kind." These exemplars are needed because they provide the "real world" examples to use when developing and testing frameworks and criteria.
3. Perform a focused literature review on the exemplar interventions, to elucidate what is currently being reported regarding evaluations of the exemplar interventions and to identify gaps in the evidence.
4. Perform an assessment of strengths and limitations of existing frameworks for assessing global health interventions.
5. Identify and/or develop of criteria to meet identified needs or gaps in existing frameworks.
6. Pilot test the draft criteria.

Assemble an Interdisciplinary Panel of Experts

Using recommendations from the global health community, our partner, and our prior work in program evaluation we were to assemble a panel of experts that included developers and implementers of interventions or programs, methods experts, and policymakers.

Select a Set of "Exemplar Interventions or Programs"

We identified a diverse set of global health interventions by considering the major causes of morbidity and mortality in developing countries or the major diseases of focus among international global health financing bodies. We developed a draft set of key dimensions for classifying global health interventions in order to map out potential exemplars in order to select a diverse set of exemplars. Having a diverse set of exemplars can help better identify generalizable strengths and weaknesses of existing or proposed frameworks and criteria, akin to the increase in generalizability that comes from enrolling a diverse set of participants in a study of the effectiveness of an intervention. TEP members were asked for input on the dimensions and preferred exemplars. TEP members were also allowed to suggest new potential exemplars that were not included in the initial set of candidates. The dimensions included the following:

- The primary population affected by the disease or the intervention—both in terms of geography as well as demographic (e.g., pregnant women, children under 5, Sub-Saharan Africa, etc.);
- Whether the intervention addresses a communicable or noncommunicable disease;
- Whether the intervention necessitates a one-time delivery or demands a sustained change in behavior or practice on the part of one or more actors;
- Whether the intervention is preventive or a treatment;
- The usual source for the delivery of the intervention—whether the health system, a nongovernmental organization (NGO) or private actors;
- A broad approximation of the level of resources required to supply the intervention (categorized as "low" or "high");
- Whether the intervention is delivered at the individual or community level (e.g., malaria bednets vs. community boreholes);
- Whether the anticipated health benefits of the intervention are at the individual or community level.

Focused Literature Review on the Exemplar Interventions for Each of the Three Chosen Global Health Exemplar Interventions

We located published systematic reviews of the effectiveness of the exemplar interventions by conducting a Medline search or using systematic reviews identified by a prior Southern California Evidence-based Practice Center project on HIV in low and middle income countries. For each of these systematic reviews, we then also retrieved the original research studies cited in them, and used both the original studies and the systematic reviews as sources of evidence when applying frameworks and developing criteria.

Assessment of Strengths and Limitations of Existing Frameworks For Assessing Global Health Interventions

We identified existing frameworks for use in public health or global health interventions by searching Medline up until March 2012, using terms such as "public health intervention," "strength of evidence," and "quality of evidence," and as well as "criteria," rating," system," etc. The vast majority of the results were reviews of the effectiveness of public health interventions (they used the terms in their texts), rather than actual systems to rate the evidence. We pulled the frameworks we were aware of through our work in evidence-based practice, contacted the TEP with what we identified, and asked for additional suggestions. We also reviewed the Web sites of the Cochrane Collaboration, the Campbell Collaboration, the SUPporting POlicy relevant Reviews and Trials (SUPPORT) project, the International Initiative for Impact Evaluation (3IE), the World Health Organizations' EVIPnet (Evidence Informed Policy Networks), SURE (Supporting the Use of Research Evidence for policy in African health systems) and the McMaster Health Forum.

Two members of the research team independently applied the six frameworks to evidence on each of the three exemplar interventions. Disagreements were settled by a group consensus process.

Identification and Development of Criteria To Meet Identified Needs or Gaps in Existing Frameworks

We presented to our TEP the results of our application of the existing public health/global health frameworks to the evidence about the global health exemplars. The TEP identified the primary gaps in existing frameworks as the lack of information about implementation and sustainability. In this project, we focused on implementation to identify existing implementation criteria (IC), we contacted the editors of the journal Implementation Science seeking their input. The editors referred us to the criteria required by their journal,[1] as well as a recent article proposing criteria for complex interventions.[2] We also reviewed the Consolidated Framework for Implementation Research,[3] which was known to us from our own implementation work and also recommended by our Task Order Officer.

We extracted all criteria from each of these three implementation criteria (IC) frameworks, and then selected those criteria that corresponded to needs we heard expressed by the TEP. We tried to include at least one criterion from each of the five domains in the Consolidated Framework for Implementation Research: intervention characteristics, outer setting, inner setting, characteristics of the individuals involved, and the process of implementation. We ended up selecting 10 implementation criteria for additional testing.

Pilot Testing of Draft Implementation Criteria

We pilot-tested these 10 implementation criteria by applying them to three articles from the evidence base of each of the three global health exemplars. For each of the 10 criteria and each of the articles, we recorded the exact text that we judged was related to the criterion, or that we could find no text in the article about a particular criterion. One reviewer extracted the data from all three articles for each exemplar and then the results for all were discussed as a group. This text for each criterion was then scored as “good,” “fair,” “poor/none,” based on the degree to which we judged the text meeting the needs of stakeholders regarding that aspect of implementation.

Results

Assemble an Interdisciplinary Panel of Experts

We recruited a panel of experts in the development and implementation of interventions, methods experts, and policymakers. We sought to include stakeholders who worked in different parts of the world, who developed or studied different types of global health problems, and who were expert in different types of evaluation methods. The members of the TEP are listed in the front matter.

Select a Set of "Exemplar Interventions or Programs"

From suggestions by experts and the literature, we selected a number of global health interventions or programs as candidates for our exemplars. We then categorized each of these using the dimensions we had developed (see Appendix A). We sent this information to our TEP and asked for the comments and input on which exemplars to choose. From this process we chose two exemplars that were on our original list—household water chlorination and interventions to prevent mother to child transmission (PMTCT) of HIV, and a third exemplar suggested by our TEP—the use of lay or community health workers (CHW) to improve maternal and child outcomes. This latter exemplar was specifically chosen since the use of chlorine to disinfect water and the use of antiretrovirals to reduce the transmission of HIV both have a "basic science" foundation. CHW is an example of intervention where this is not the case. Table 1 displays how we classified each exemplar according to our domains of diversity.

Table 1. Diversity of exemplar interventions across key criteria

Dimension/Examples	PMTCT	Household Water Chlorination	CHW
Populations affected by a disease – maternal health, adults, elderly, children	pregnant women	primarily children under 5 (diarrhea)	all, primarily children under 5
Populations affected by a disease – Africa vs. Asia vs. Latin America	Sub-Saharan Africa, Southeast Asia	all	all, primarily children under 5
Communicable vs. non-communicable disease	communicable	communicable	both
One time vs. continuous behavior	time limited	continuous	depends
Preventive or treatment	preventive	preventive	preventive and treatment
Delivery source for intervention	often publicly provided by NGOs	could be either	generally publicly provided by health system
Resource dimension	high	low	unknown
Individual vs. Community - level of impact	mostly individual	both	mostly individual
Individual vs. Community - level of intervention	individual	individual	community
Urban vs. Rural	urban and rural	urban and rural	mostly rural?

Focused Literature Review on the Three Chosen Global Health Exemplar Interventions

We identified the following systematic reviews and original research studies for each intervention:

Household Water Chlorination

The evidence base for household water chlorination stemmed from two 2007 systematic reviews of randomized and nonrandomized studies; one review[4] was specific to household chlorine use, and the other[5] assessed numerous water quality interventions and for these we isolated the chlorination studies where possible. Both reviews broadly searched numerous databases including Medline, Embase, LILACS and Cochrane. The review specifically about household chlorine use did not restrict based on study design, while the other reviews included only RCTs and quasi-randomized controlled trials. Three RCTs were included in both analyses. Meta-analyses of beneficial outcomes included 10 and 5 studies of chlorine, respectively. We applied the evidence frameworks to outcomes of both water quality and diarrheal morbidity as both were considered in the systematic reviews and due to the self-reported nature of diarrheal episodes in these studies. The pooled analysis in the review specific to household chlorine use found a .71 relative risk of child diarrhea outcomes (95% confidence interval [CI]: 0.58–0.87) and a 0.20 relative risk of water contamination outcomes (95% CI: 0.13–0.30) due to the interventions. Implementation data came from review of three of the original studies included in these reviews.[6-8] Studies came from urban and rural settings within South America, Africa, Southeast and Central Asia.

PMTCT

The primary piece of evidence on PMTCT was a high-quality Cochrane systematic review updated in 2011.[9] The review focused on efficacy and safety of various antiretroviral regimens in low income countries; the primary outcome was percent of infants HIV positive at various time points postpartum. The review included only RCTs; there were 25 including almost 19,000 participants total. The review reports on the efficacy of various regimens in breast-feeding and non breast-feeding mothers and concludes that triple antiretroviral therapy is the most effective regimen. All antiretroviral interventions were found safe in the short term.

This review was complemented by a 2011 systematic review on cost-effectiveness of PMTCT in low and middle income countries.[10] Nineteen peer-reviewed cost-effectiveness studies were included; 15 focused on antiretrovirals (ARV), while others included primary prevention of HIV, HIV testing strategies, prevention of unwanted pregnancies, and replacing of breast-feeding with other means. All articles conducted modeling with hypothetical cohorts; ARV regimens and their efficacy rates were often taken from the same trials included in the Cochrane review.[9] Sixteen of the 19 studies concluded that PMTCT is cost-effective. Two others involved HIV testing in low prevalence areas and one other was conducted when the cost of antiretrovirals was significantly higher than in recent times.

The third review included only studies of ARV conducted in sub-Saharan Africa.[11] It was limited to peer-reviewed RCTs; meta-analysis of 10 resulted in an estimate of about 50 percent efficacy in reducing HIV transmission from mother to child.

Lay or Community Health Workers

For lay workers we identified a Cochrane Review on "lay health workers in primary and community health care for maternal and child health and the management of infectious diseases" that was updated as of October 2009.[12] This review searched Medline, Embase, CINAHL, Cochrane and a number of other databases through early 2009. Eligible studies had to be RCTs. A "lay health worker" was defined as "any health worker carrying out functions related to health care delivery, trained in some way in the context of the intervention, and having no formal or professional or paraprofessional certificate or tertiary education degree." A large number of different studies and outcomes were included in this review (82 studies and 25 different outcomes; of course most studies only assessed a limited number of outcomes). Assessed outcomes included uptake of immunizations, promoting breastfeeding, mortality/morbidity among children less than 5 years of age, neonatal mortality, reported illnesses in children, care-seeking behavior, and various tuberculosis outcomes. From these we chose the outcome "reduce morbidity in children under 5 years of age" as it seemed to be both an outcome very important to communities and it had enough studies to make a meta-analysis meaningful (3 trials). Seven trials were included, which were conducted in Bangladesh (2 studies), Thailand, Burkina Faso, India, Nepal, and Vietnam. The pooled analysis of all 7 studies was a reduction in reported illness in children of 0.86 (95% CI 0.75, 0.99). The conclusion of the review was that the use of lay health workers, compared with usual care "may lead to slightly fewer children who suffer from fever, diarrhea and pneumonia."

Assessment of Strengths and Limitations of Existing Frameworks For Assessing Global Health Interventions

In our efforts to identify frameworks, we reviewed more than 50 documents identified by our search methods described earlier. Many contained helpful suggestions of areas to focus on when assessing public health interventions in low and middle income countries. However, only six included clear criteria with a rating system and instructions for application to a body of evidence on an intervention. Most of these frameworks had some areas of overlap, such as greater weight given to evidence generated by randomized controlled trials. However, the relative weight given to other study designs, as well as the consideration of factors such as details of implementation and the generalizability of findings received differing amounts of consideration across the frameworks. We briefly summarize main points of these six frameworks in Table 2.

Table 2. Public health/global health frameworks

Framework Name	Grades Assigned	Domains for Grading Evidence	Notes on Domains
Tang, Choi, Beaglehole *Journal of Epidemiology of Community Health 2008*[13] *(*Drawn from Tang et al.'s Table 1 for grading of evidence on association, repeatability and causal mechanism)	Grades of 1 (strong), 2 (weak), and 3 (insufficient). Expanded categories include grade 2A (probable), 2B (possible) and 2C (limited).	Association	"High" association is defined as a RR of greater than 2. Otherwise "low" or "none."
		Repeatability	Wide or Limited
		How it works	How it works is known or not known
Grading of Recommendations Assessment, Development and Evaluation (GRADE)[14] (Summarized from BMJ article Table 1.)	Four grades assigned: high, moderate, low, and very low quality of evidence. For "high" quality studies, definition is "We are very confident that the true effect lies close to that of the estimate of the effect."	Randomized trials start with a "high" initial quality grade, observational studies start with a "low" grade.	Grades can be moved down depending on factors such as risk of bias or inconsistency, or up in light of a large measured effect or evidence of a dose-response.
Highest Attainable STandard of Evidence (HASTE)[15]	Four grades assigned: Strong (Grade 1), Conditional (Grade 2), Insufficient (Grade 3), and Inappropriate (Grade 4). Conditional (Grade 2) has subcategories of Probable, Possible and Pending.	Efficacy	Whether consistent, limited or inconsistent
		Biological plausibility	
		Implementation data availability	Whether available or not
U.S. Community Preventive Services Task Force (USCPSTF), Briss, Saza, Pappaioanou et al. 2000[16]	Evidence is characterized as strong, sufficient or insufficient	Execution	Good or Fair
		Design suitability	Greatest (RCTs), Moderate (no concurrent comparison group) or Least
		Number of studies	
		Consistent	"Generally consistent in direction and size"
		Effect size	Sufficient or large, defined on a case-by-case basis based on Task Force opinion
		Expert opinion	Whether used or not
Australian National Health and Medical Research Council (NHMRC) 2009[17]	Four grades assigned: A=Excellent, B=Good, C=Satisfactory, D=Poor. Grades of "A" can be trusted to guide practice; grades of D concludes the body of evidence is weak and recommendation must be applied with caution.	Evidence base	'Evidence hierarchy' places systematic reviews of RCTs with "low risk of bias" highest
		Consistency of evidence	
		Clinical impact	very large, substantial, moderate, slight
		Generalizability	Highest grade awarded if "populations studies in body of evidence are *the same* as the target population for the guideline" (emphasis added)
		Applicability	

Table 2. Public health/global health frameworks (continued)

Framework Name	Grades Assigned	Domains for Grading Evidence	Notes on Domains
UK National Health Service (NHS) Health Development Agency 2005[18]	Four grades assigned: A, B, C and D.	Efficacy	High quality meta-analyses and systematic reviews of RCTs with very low risk of bias rated highest level of evidence.
		Evidence of corroboration	Strong evidence of corroboration defined as "Consistent findings in two or more studies of ++ quality carried out within the UK and applicable to the target population, providing evidence on salience and implementation." ++ is defined as is efficacy above.

Table 3 below summarizes our findings from the application of the six existing evidence frameworks to the three global health exemplars. Household water chlorination had two potential primary outcome measures of interest with which we undertook this exercise: water quality (often objectively measured via water testing) and diarrheal outcomes, which are predominantly self-reported but are often the primary outcomes reported in trials of chlorination studies. Since there was no guidance in the frameworks themselves regarding what types of outcomes are appropriate for their application, we adopted both measures. For PMTCT reviews, we adopted the convention of assessing its body of evidence according to the rate of HIV infection in children assessed at various times during the first year of life. For the literature on Lay or Community Health Workers (CHWs), we adopted the literature's convention of assessing its effectiveness according to the outcome of under-5 morbidity and mortality from any cause.

Table 3. Results on exemplars applied to six evidence frameworks

Outcomes	Tang et al.[13]	GRADE[14]	HASTE[15]	Community Preventive Services Task Force[16]	Australian NHMRC[17]	NHS Health Development Agency[18]
Household Water Chlorination						
Water Quality	Grade 1 level 1 Strong	⊕⊕⊕⊕ High quality of evidence	Grade 1-Strong	Strong	"B" - Good	"A"
Diarrhea	Grade 2b level 1 possible	⊕⊕⊕ Moderate quality of evidence	Grade 2b - Possible	Strong	"C" - Satisfactory	"B"
Preventing Mother-to-Child Transmission, All Regimens Included						
HIV infection in child within year of birth	Grade 2b level 1 possible^	⊕⊕⊕⊕ High quality of evidence	Grade 1-Strong	Strong	"A" - Excellent	"A"
Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared with usual care						
Mortality and Morbidity	Grade 2b Level 2 Possible*	⊕⊕ Low quality of evidence	Grade 3 - Insufficient	Strong	"B" - Good	"C"

*Grade 2c Level 2 if repeatability outside Southeast Asia is not considered acceptable.
^Tang et al. grade for PMTCT is due to strict rule that only interventions with relative risk (RR)>2 qualify as "strong" evidence. If this rule is flexible we would rate PMTCT as "Grade 1 Level 1 Strong" by Tang et al. categorizations.

We focus our attention to the frameworks themselves versus the precise grades assigned to the various exemplars. Operationalizing these frameworks involved a good deal of subjective reasoning and more details for the precise grades assigned to a particular outcome are available in Appendix B.

When measured water quality is the primary outcome from studies of household water chlorination, we generally find that the six frameworks assign a high grade to evidence of its effectiveness. Five of the six frameworks assign their highest possible grade; the Australian National Health and Medical Research Council (NHMRC) framework assigns a grade of "B" (Good). However, when the self-reported diarrheal outcomes are considered, the evidence frameworks generally conclude that the evidence is weak or moderate quality. Only the Community Preventive Services Task Force (CPSTF) framework continues to assign household water chlorination its highest grade for evidence ("Strong"). All of the remaining frameworks downgrade the evidence base by at least one grade in their categorizations, with the evidence classifications now ranging from the highest categorization of "Strong" by the CPSTF framework, to the next-to-lowest grade of "C – Satisfactory" within the Australian NHMRC framework.

For interventions involving community or lay health workers to prevent under-5 morbidity and mortality compared with usual care, the various frameworks generally rate the evidence as being of low or moderate quality with the exception of CPSTF which assigns the highest grade of "Strong." HASTE, on the other hand, would rate this same body of evidence as grade 3 "Insufficient," and GRADE also assigns it a "⊕⊕ Low quality of evidence."

For PMTCT studies, all of the frameworks assign their highest possible grade to the body of evidence with the exception of the framework by Tang et al.,[13] which assigns a "Grade 2B, Level 1 Possible." However, this grade is the result of a strict interpretation of the rule that only interventions with a relative risk (RR) of greater than 2 qualify as "strong." If there is some flexibility with this strict cutoff both reviewers assigned this exemplar agree its rating would change to the highest grade of "Grade 1 Level 1 Strong."

Therefore, for two of the four interventions / outcomes (chlorine use and water quality, PMTCT and HIV infection in children within 1 year) there was very good agreement across the frameworks that these interventions were at or near the highest rating on the scale, whereas for the other two interventions / outcomes at least one framework classified the evidence at the top of the rating scale while one or more of the other frameworks classified the evidence at the mid-point or even at the bottom of the rating scale.

In Table 4 we make direct comparisons across the six evidence frameworks according to eight key criteria that we considered to be important for our context of informing community or scale-level policy for global health interventions or programs in developing country settings. These criteria draw from the frameworks themselves as well as considerations by the project team that came about through this exercise. They include (1) how strict or explicit the rules are for classifying the strength of evidence; (2) the magnitude of potential benefits versus harms; (3) what role, if any, context is taken into consideration in evaluating the evidence; (4) how much is known about the details of implementation; (5) whether the ease of implementing the intervention or program is taken into consideration; (6) total costs for the program or intervention; (7) sustainability of the program or intervention, both cost-wise and programmatically; and (8) the ease of operationalizing these frameworks in practice (as, presumably or hopefully, a policymaker may need/want to do).

None of the frameworks takes into consideration the costs or sustainability of the intervention in assigning its grades to the evidence (Note: GRADE has a separate set of guidance statements for developing recommendations, and include cost). While some frameworks have explicit criteria for assigning grades on strength of evidence,[14,16-18] others require more individual interpretation.[13,15] Frameworks also differ in whether or not they assess magnitudes of benefits versus harms, and whether or not context is taken into consideration. Similarly, wide differences in whether and the ways in which details of implementation are taken into account exist across frameworks. HASTE calls for a detailed assessment of implementation data while CPSTF, Australian NHMRC, Tang et al. and GRADE do not mention implementation

Table 4. Comparison of six frameworks for assessing the evidence about global health interventions

Frameworks	Domains							
	Assessment of Strength of Evidence How Explicit? How Strict?	**Assessment Of Magnitude Of Benefit vs. Harm**	**Assessment of Context**	**Assessment of Implementation How Much is Known?**	**Assessment of Implementation How Easy is Implementation?**	**Costs**	**Sustainability**	**Comments/Ease of Application**
Tang et al.[13]	Not as explicit as some others, not as strict as GRADE	Explicit assessment of magnitude of benefit, no explicit assessment of harm	Potentially assessed as part of "widely demonstrated"	Not directly assessed	Could potentially be included in repeatability	Not assessed	Not assessed	Relatively easy to apply, but operationalizing "repeatability" is a challenge
GRADE[14]	Very explicit, strictest criteria	Explicit assessment	Not assessed	Not assessed	Not assessed	Not assessed*	Not assessed	Easy to apply, widely used and recognized framework
HASTE[15]	Not explicitly assessed, more lenient criteria than GRADE	Not explicitly assessed	Not assessed	Detailed assessment of implementation science data	No explicit assessment of ease of implementation but could readily be incorporated	Not assessed	Not assessed	Relatively easy to apply, some challenge with categories (for example, the situation where efficacy is consistent, biologically plausible and implementation data scarce has no category). The published example contains far more data collection on implementation than what would normally be feasible

Table 4. Comparison of six frameworks for assessing the evidence about global health interventions (continued)

Frameworks	Domains							
	Assessment of Strength of Evidence How Explicit? How Strict?	Assessment Of Magnitude Of Benefit vs. Harm	Assessment of Context	Assessment of Implementation How Much is Known?	Assessment of Implementation How Easy is Implementation?	Costs	Sustainability	Comments/ease of Application
Community Preventive Services Task Force[16]	Explicitly assessed, more lenient criteria than GRADE	Explicit assessment of benefits with expert judgment for threshold, no explicit assessment of harms	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Easy to apply
Australian NHMRC[17]	Explicitly assessed, more lenient criteria than GRADE	Explicit assessment	Yes	Not directly assessed, could potentially be included in applicability	Not directly assessed, could potentially be included in applicability	Not assessed, could potentially be included in applicability.	Not directly assessed, could potentially be included in applicability	Very difficult to apply. Potential is large for poor inter-rater reliability.
NHS Health Development Agency[18]	Explicitly assessed, criteria approximately as strict as GRADE	Not explicitly assessed	Not directly assessed Indirectly included in corroboration	Included as a component of corroboration	Not directly assessed, but potentially could be part of corroboration	Not assessed	Not assessed	Most difficult to apply. Potential is large for poor inter-rate reliability

*Costs are considered in GRADE's factors to consider when making a recommendation[19]

Identification and Development of Criteria To Meet Identified Needs or Gaps in Existing Frameworks

The research team presented initial findings of this exercise via teleconference to members of the TEP for feedback on what elements were missing that might be necessary to inform community or scale-level global health policy and programming.

During this discussion, a number of issues arose, including challenges to describing the variability in the intervention (what constitutes household water chlorination is different in terms of variability that what constitutes a lay health care worker), the importance of context when assessing effectiveness of interventions, and the distinction between what is the intervention and what is the implementation. While issues of study design were brought up and discussed, (e.g. the role of randomized studies as opposed to nonrandomized studies) input from the TEP indicated that better information about implementation is needed, regardless of the study design, rather than trying to further argue the merits of randomization. Thus, we determined that the next step for this project was to identify or develop implementation criteria that could be used to supplement any existing frameworks.

We identified three existing sets of criteria for reporting implementation information, from the editors of the journal Implementation Science,[1] the Consolidated Framework for Implementation Research[3] and the proposed criteria for reporting the development and evaluation of complex interventions in health care (CReDECI).[2] Using the TEP-identified issues as a guide, we selected from these the following 10 criteria as potentially relevant to report for implementation of global health interventions, and therefore worth testing in a pilot study. The examples in Table 5 below are ones we developed or adapted from existing examples in the original articles.

Table 5. Global framework—rating implementation criteria

Criterion	Example
Criterion #1 - Intervention Characteristics: Intervention/Program source (From CFIR, Damschroder, 2009)[3]	Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.
Criterion #2 - Intervention Characteristics: A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[1,2]	The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries.

Table 5. Global framework—rating implementation criteria (continued)

Criterion	Example
Criterion #3 - Intervention Characteristics: Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[1,2]	Explanation/Example: "Our preliminary qualitative field work showed that individual behaviours were influenced by collective behaviours and social norms, and sustained by a complex, multilevel network of relationships within the community. We therefore developed a multilevel strategy targeting: community stakeholders, newborn stake holders, and households with immediate support groups. At each level, the target group consisted of individuals who were identified to have key roles as influencers, decision makers, supporters, and practitioners of newborn care and normative behaviour within the community. The support of community stake holders such as village heads, community leaders, respected members, priests, and teachers was crucial in building trust with the community and ensuring acceptance of the programme. The newborn stakeholder target group included traditional newborn-care providers and birth attendants, unqualified medical practitioners, and, to a lesser extent, health system workers, some of whom had strategic access to the newborn and mother during post-partum confinement, were perceived by the community as domain experts, and played an active part in sustaining targeted practices. Health system workers such as auxillary nurse midwives were engaged only at the community level as part of newborn stakeholder group meetings in order to keep contamination of the intervention into control clusters to a minimum. The household target group included the pregnant woman or mother, who was the primary care provider, but usually not empowered to make decisions; the mother-in-law, who was usually the key decision maker on newborn-care practices; other female members who played supportive roles; and male members, including the father-in-law and husband, who controlled access to the household, made financial and logistical arrangements, and influenced care-seeking decisions. The family's immediate support group included neighbours and relatives who influenced family behaviours and helped with deliveries."[23]

Table 5. Global framework—rating implementation criteria (continued)

Criterion	Example
Criterion #4 - Outer Setting: External policies and incentives (From CFIR, Damschroder, 2009)[3]	How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.
Criterion #5 - Intervention Characteristics: Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[1] **The detailed description should include:** a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker) b. Characteristics of the recipients c. The setting d. The mode of delivery (such as face-to-face) e. The intensity of the intervention/program (such as the contact time with participants) f. The duration (such as the number of sessions and their spacing interval over a given period) g. Adherence or fidelity to delivery protocols h. A detailed description of the intervention/program content provided to each study group	
Criterion #6 - Intervention Characteristics: Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3]	The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.
Criterion #7 - Population needs (From CFIR, Damschroder, 2009)[3]	The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.
Criterion #8 - Process of implementation: Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[1,2]	"The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."[2]
Criterion #9 - Description of materials: Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[2]	"The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroom based and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."[23]

Table 5. Global framework—rating implementation criteria (continued)

Criterion	Example
Criterion #10 - Process of Implementation: Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[2]	Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."[2]

Pilot Test Draft Criterion

For the pilot testing of the implementation criteria, Tables 6 and 7 present summary findings for household water chlorination, lay health workers, and prevention of mother-to-child-transmission of HIV, respectively, by criterion and by article. More detailed tables assessing what text was found and how we judged it meeting the criteria can be found in Appendix C. Two findings from the summary tables are worth noting: first, the reporting of implementation information is highly variable both within and across articles, with some articles reporting a great deal of information about some criteria and almost nothing about others, and likewise some articles reporting a great deal about most criteria while other articles report almost nothing about most criteria; and second, some criteria were more difficult to judge than others. The criterion on the outer setting, in particular, was one we had a hard time converting into an operational measure. Several of the criteria seemed to be somewhat overlapping, suggesting the possibility that these could be consolidated, for example criterion 5h "A detailed description of the intervention/program content provided to each study group" and criterion 9 "Description of all materials or tools used for the implementation."

Table 6. Pilot test results of applying draft implementation reporting criteria to published studies of assessments of the effectiveness of three Global Health Intervention exemplars

Criteria	3 Studies of Household Water Chlorination (Number of Articles Reporting Criterion)			3 Studies of Preventing Mother-to-Child Transmission (Number of Articles Reporting Criterion)			3 Studies of Lay/Community Health Workers (Number of Articles Reporting Criterion)			Overall (Number of Articles Reporting Criterion)		
	Good	**Fair**	**Poor/None**	**Good**	**Fair**	**Poor/None**	**Good**	**Fair**	**Poor/None**	**Good**	**Fair**	**Poor/None**
1.	1	2	0	2	0	1	1	2	0	4	3	2
2.	0	1	1	0	3	0	2	0	1	2	4	3
3.	0	0	3	0	2	1	1	0	2	1	2	6
4.	0	0	3	0	1	2	0	0	3	0	1	8
5a.	1	1	1	1	0	2	1	2	0	3	3	3
5b.	3	0	0	2	1	0	2	1	0	6	1	2
5c.	3	0	0	2	1	0	3	0	0	8	1	0
5d.	3	0	0	2	1	0	1	1	0	6	3	0
5e.	2	0	1	0	2	1	1	0	2	3	2	4
5f.	2	1	0	1	1	1	1	0	2	4	2	3
5g.	0	2	1	0	1	2	2	1	0	2	4	3
5h.	0	1	2	2	1	0	1	1	2	3	3	4
6.	0	1	2	0	0	3	0	0	3	0	1	8
7.	0	1	2	1	0	2	1	0	2	2	1	6
8.	0	0	3	2	0	1	1	0	2	3	0	6
9.	0	0	3	1	1	1	1	1	1	2	2	5
10.	0	0	3	0	2	1	1	0	2	1	2	6
TOTALS	**15**	**10**	**25**	**16**	**17**	**18**	**20**	**9**	**22**	**50**	**35**	**69**

Table 7. Number of criteria met, by article

Article	Good	Fair	Poor/None
Household Water Chlorination			
Quick et al[7]	4	5	7
Luby et al [8]	6	2	9
Crump et al[6]	5	3	8
Preventing Mother-to-Child Transmission			
Kesho Bora Study Group[20]	6	7	4
Lussiana et al[21]	0	5	12
Kim et al[22]	9	6	2
Lay Health Workers			
Kumar et al[23]	12	1	4
Koyate et al[24]	2	5	9
Sloan et al[25]	6	3	7

Discussion

The principal results of this project are:

- Existing frameworks for assessing global health interventions or programs when applied to the same body of evidence for three exemplar interventions yielded somewhat to markedly different assessments of the strength of evidence.
- The most important gap in existing frameworks for assessing global health interventions is the lack of information about implementation, including costs, sufficient for policymakers to make judgments about the sensitivity of effectiveness to differences in context, and the scalability and sustainability of the intervention.
- A pilot test of established criteria for reporting implementation data showed that existing publications of effectiveness of global health interventions vary greatly in the amount of implementation data reported, with more than 40 percent of criteria having poor or absent reporting, and another 20 percent having only fair reporting.

Regarding the first principal result, we also note that even within the same framework, different team members often initially reached different conclusions, largely due to a wide scope for individual interpretation that raises concerns about inter-rater reliability. Also, we note that whereas almost all existing frameworks have explicit assessments of strength of evidence based in some part on study design, almost no frameworks have any assessments of the role of context or costs.

Regarding the third principal result, we note that, like the global health strength of evidence frameworks, the challenge in the operability of many of the criteria raises concern about inter-rater reliability. Additionally, the overlapping nature of some of these criteria indicates that additional work is needed to define what constitutes adequate reporting of implementation details sufficient for stakeholders to make assessments about the sensitivity of effectiveness to context, scalability, and sustainability.

In our judgment, these three results inevitably lead to the conclusion that no existing framework is sufficiently useful for stakeholders to make the kinds of decisions they need to make about global health interventions. Therefore, more work is needed to either adapt one or more existing frameworks, or develop an entirely new framework.

In addition to rating the strength of evidence, (which GRADE refers to as quality of evidence) GRADE has advice for developing recommendations.[26] For recommendations, GRADE lists four key factors: the balance between the desirable and undesirable consequences of the alternatives; the quality of the evidence about the desirable and undesirable consequences; uncertainty or variability in values and preferences; and cost. This advice could also be applied to global health decision making. To date, GRADE has produced explicit advice for how to assess resource use;[27] and stated that the specific context is critical for considering resource use.[19]

A recent series of papers from the Task Force on Developing Health Systems Guidance took this advice one step further.[28] The Task Force lists a number of factors that may inform decisions about the strength of recommendations regarding policy options, some of which are adopted from GRADE and some of which are in addition to GRADE:

- Whether there is uncertainty about the balance of benefits versus harms and burdens
- The quality of the evidence from the systematic review (very low, low, moderate, high)
- Whether there is uncertainty or variability in values and preferences among stakeholders

- Whether there is uncertainty about whether the net benefits are worth the costs or about resource use
- Whether there is uncertainty about the feasibility of the intervention (or about local factors that influence the translation of evidence into practice, including equity issues)
- Ease of implementation at the systems level, including governance arrangements (e.g., changes needed in regulations), financial arrangements (e.g., the extent to which the options fit with financing models within settings), and implementation strategies (e.g., how to provide the skills and experience needed among implementers or facilitators)
- Socio-political considerations, e.g., how the proposed options relate to existing policies, values within the political system in relation to issues such as equity or privatization, and economic changes

Our project indicates that continuing to develop these factors, particularly about implementation, would be beneficial.

Limitations

There are some notable limitations of this study. First, we might have missed some global health frameworks in our identification process. Second, not all of these frameworks necessarily had as goals the assessment of information on costs, or contexts, or implementation. Nevertheless, these are crucial aspects of the assessment of evidence about global health interventions, and therefore even if this was not an explicit goal of an existing framework then it is important to note its failure to do so. Third, we only assessed the performance of the existing frameworks with three exemplars, and they may have performed differently on other global health interventions. However, we did purposefully pick our exemplars to represent a diversity of interventions, and furthermore note that our assessment of the frameworks on three exemplars is in general two greater in number than the assessments that were part of the original framework presentation, which usually demonstrated its use with a single example. Furthermore, evaluating the frameworks on additional exemplars is unlikely to change the existing conclusions of variability both within and across frameworks in how evidence is assessed, and will not change the identification of context, costs, and implementation data as important missing domains of these frameworks. Fourth, we relied heavily on input from our technical expert panel, and therefore our results may have been sensitive to the actual participants on our panel and at each stage of the process. Further evaluation of these results with a wider group of stakeholders is warranted. However, these stakeholders' identification of a need for more data about implementation is consistent with the increasing recognition of the importance of implementation reporting in other health-related fields. Fifth, we selected the implementation criteria to assess based on our own judgment, and this would benefit from stakeholder input. Finally, when making judgments about the adequacy of reporting of implementation information we used our own assessments, and this would be better made using the assessments of actual stakeholders, in terms of the adequacy of the reported information to guide decisions about scalability and sustainability.

Conclusions and Future Research Needs

Existing frameworks for evaluating evidence for global health interventions are both too variable to have confidence in their application and do not meet the needs of stakeholders. Additional work needs to be done to develop a framework that will meet stakeholders' needs, and a focus on elements of the reporting of implementation data are a crucial need for a framework to be useful. This will involve both the identification of implementation criteria stakeholders judge to be useful, and then the assessment of whether or not those criteria change or improve decisionmaking. The experience of reporting guidelines for clinical trials offers an example of how a process of evidence and expert opinion can create tools that help the design and reporting of studies of medical interventions.[29] Better conducted and reported clinical trials then lead to more informed decisionmaking when developing clinical practice guidelines. The development of reporting guidelines for global health interventions, although a formidable task, has the potential to similarly improve global health decisionmaking.

The reporting of adequate information on implementation will require more text than authors and peer-reviewed journals are currently devoting to the methods section of their publications. It is not coincidental that the article that had the highest number of "good" scores for implementation criteria devoted 1,800 words to describing the intervention and the implementation, which is 50 percent to 80 percent of the word limit many leading journals place on the entire publication. Making additional descriptive material available on the Internet may be one solution to this restriction.

References

1. Michie S, Fixsen D, Grimshaw JM, et al. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci. 2009;4:40. PMID 19607700.
2. Mohler R, Bartoszek G, Kopke S, et al. Proposed criteria for reporting the development and evaluation of complex interventions in healthcare (CReDECI): guideline development. Int J Nurs Stud. 2012 Jan;49(1):40-6. PMID 21924424.
3. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:50. PMID 19664226.
4. Arnold BF, Colford JM, Jr. Treating water with chlorine at point-of-use to improve water quality and reduce child diarrhea in developing countries: a systematic review and meta-analysis. Am J Trop Med Hyg. 2007 Feb;76(2):354-64. PMID 17297049.
5. Clasen T, Schmidt WP, Rabie T, et al. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. BMJ. 2007 Apr 14;334(7597):782. PMID 17353208.
6. Crump JA, Otieno PO, Slutsker L, et al. Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhoea in areas with turbid source water in rural western Kenya: cluster randomised controlled trial. BMJ. 2005 Sep 3;331(7515):478. PMID 16046440.
7. Quick RE, Venczel LV, Mintz ED, et al. Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. Epidemiol Infect. 1999 Feb;122(1):83-90. PMID 10098789.
8. Luby SP, Agboatwalla M, Painter J, et al. Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomised controlled trial. Trop Med Int Health. 2006 Apr;11(4):479-89. PMID 16553931.
9. Siegfried N, van der Merwe L, Brocklehurst P, et al. Antiretrovirals for reducing the risk of mother-to-child transmission of HIV infection. Cochrane Database Syst Rev. 2011(7):CD003510. PMID 21735394.
10. Johri M, Ako-Arrey D. The cost-effectiveness of preventing mother-to-child transmission of HIV in low- and middle-income countries: systematic review. Cost Eff Resour Alloc. 2011;9:3. PMID 21306625.
11. Chigwedere P, Seage GR, Lee TH, et al. Efficacy of antiretroviral drugs in reducing mother-to-child transmission of HIV in Africa: a meta-analysis of published clinical trials. AIDS Res Hum Retroviruses. 2008 Jun;24(6):827-37. PMID 18544018.
12. Lewin S, Munabi-Babigumira S, Glenton C, et al. Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases. Cochrane Database Syst Rev. 2010(3):CD004015. PMID 20238326.
13. Tang KC, Choi BC, Beaglehole R. Grading of evidence of the effectiveness of health promotion interventions. J Epidemiol Community Health. 2008 Sep;62(9):832-4. PMID 18701736.
14. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924-6. PMID 18436948.
15. Baral S, Wirtz A, Sifakis F, et al. The Highest Attainable Standard of Evidence (HASTE) for HIV/AIDS interventions: towards a public health approach to defining evidence. Public Health Rep. 2012 Nov;127(6):572-84. PMID 23115382
16. Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based Guide to Community Preventive Services—methods. The Task Force on Community Preventive Services. Am J Prev Med. 2000 Jan;18(1 Suppl):35-43. PMID 10806978.
17. National Health and Medical Research Council. NHMRC levels of evidence and grades for recommendations for developers of guidelines. Australia: December 2009.

18. Wightman A, Ellis S, Cullum A, Sander L, Turley R. Grading evidence and recommendations for public health interventions: developing and piloting a framework. Health Development Agency. 2005. http://www.nice.org.uk/nicemedia/docs/grading_evidence.pdf . Accessed December 2012.

19. Guyatt GH, Oxman AD, Kunz R, et al. Incorporating considerations of resources use into grading recommendations. BMJ. 2008 May 24;336(7654):1170-3. PMID 18497416.

20. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery and breastfeeding for prevention of mother-to-child transmission of HIV-1: the Kesho Bora Multicentre Collaborative Study rationale, design, and implementation challenges. Contemp Clin Trials. 2011 Jan;32(1):74-85. PMID 20854932.

21. Lussiana C, Clemente SV, Ghelardi A, et al. Effectiveness of a prevention of mother-to-child HIV transmission programme in an urban hospital in Angola. PLoS One. 2012;7(4):e36381. PMID 22558455.

22. Kim MH, Ahmed S, Buck WC, et al. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.

23. Kumar V, Mohanty S, Kumar A, et al. Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. Lancet. 2008 Sep 27;372(9644):1151-62. PMID 18926277.

24. Kouyate B, Some F, Jahn A, et al. Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial. Malar J. 2008;7:50. PMID 18364043.

25. Sloan NL, Ahmed S, Mitra SN, et al. Community-based kangaroo mother care to prevent neonatal and infant mortality: a randomized, controlled cluster trial. Pediatrics. 2008 May;121(5):e1047-59. PMID 18450847.

26. Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. BMJ. 2008 May 10;336(7652):1049-51. PMID 18467413.

27. Brunetti M, Shemilt I, Pregno S, et al. Grade guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol. 2012 Aug 3. PMID 22863410.

28. Lewin S, Bosch-Capblanch X, Oliver S, et al. Guidance for evidence-informed policies about health systems: assessing how much confidence to place in the research evidence. PLoS Med. 2012 Mar;9(3):e1001187. PMID 22448147.

29. Consort Statement. Consort Transparent Reporting of Trials, Available at: www.consort-statement.org/. 2010. Accessed December 2012.

Abbreviations

3IE	International Initiative for Impact Evaluation
AHRQ	Agency for Healthcare Research and Quality
ARV	antiretrovirals
CHW	Community Health Worker
CPSTF	Community Preventive Services Task Force
EPC	Evidence-based Practice Center
EVIPnet	Evidence Informed Policy Networks
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HASTE	Highest Attainable Standard of Evidence
HDA	Health Development Agency
IC	Implementation Criteria
NHMRC	National Health and Medical Research Council
PMTCT	Prevent Mother-to-Child Transmission
UK NHS	National Health Service
RCTs	Randomized Clinical Trials
SUPPORT	SUPporting Policy relevant Reviews and Trials
SURE	Supporting the Use of Research Evidence
TEP	Technical Expert Panel
USAID	U.S. Agency for International Development

Appendix A. Overview of Potential Exemplars

Classification Dimensions

Dimension/Examples	Bednets for Malaria	Malaria Vaccine	MTCT for HIV	Household Water Chlorination	Condom Distribution	HIV RX as Prevention	Oral Rehydration Salts (ORS)	Handwashing	ARI/ Improved Cook Stove Adoption	Source Water Improvement
Populations affected by a disease										
- maternal health, adults, elderly, children	all	all	pregnant women	primarily children under 5 (diarrhea)	adults	people with/without HIV	primarily children under 5 (diarrhea)	all, predominantly children (diarrhea)	all, predominantly women and children	all, predominantly children
- Africa vs. Asia vs. Latin America	all	all	Sub-Saharan Africa, Southeast Asia	all	all	all	all	all	all	all
Communicable vs. non-communicable disease	communicable	communicable	communicable	communicable	communicable	Communicable & non-communicable	communicable	communicable	Non-communicable	communicable
One time vs. continuous behavior	continuous	one time	time limited	continuous	one time per episode	continuous	one time per episode	continuous	both (initial adoption and continued use)	one time installation; continuous on part of households to use
Preventative or treatment	prevent	prevent	prevent	prevent	prevent	prevent / treatment	treatment	prevent	prevent	prevent
Delivery source for intervention	either	publicly provided	often publicly provided by NGOs	could be either	could be either	often publicly provided by NGOs	could be either	private decisions	private	public

Dimension/Examples	Bednets for Malaria	Malaria Vaccine	MTCT for HIV	Household Water Chlorination	Condom Distribution	HIV RX as Prevention	Oral Rehydration Salts (ORS)	Handwashing	ARI/ Improved Cook Stove Adoption	Source Water Improvement
Resource dimension	low	?	high	low	low	high	low	low	high	High
Individual vs. Community - level of impact	mostly individual	mostly individual	mostly individual	both	mostly individual	both	mostly individual	both	individual	both
Individual vs. Community - level of intervention	individual	individual	individual	individual	individual	individual/population	individual	individual	individual	community
Urban vs. Rural	rural	rural	urban and rural	urban and rural	urban and rural	urban and rural	urban and rural	urban and rural	both	both, mostly rural

Appendix B. Using Six Different Frameworks To Assess the Evidence for Three Examples of Health Interventions or Programs

Tang Framework[1]

Grade 1 (strong)

What works is known: a high association is shown; and repeatability is widely demonstrated

Level 1 how it works is known

Level 2 how it works is not known

Grade 2 (weak)

What works is known: but a low association is shown; and/or repeatability is limited

Level 1 how it works is known

Level 2 how it works is not known

Grade 3 (insufficient)

What works is not known: evidence is insufficient

Table B.1 Grading of evidence based on association, repeatability and causal mechanism

3–Grades	3–Grades expanded	Association	Repeatability	How it works
1 (strong)	1 (strong)	High	Wide	Known
				Not known
	2A (probable)	High	Limited	Known
				Not known
2 (weak)	2B (possible)	Low	Wide	Known
				Not known
	2C (limited)	Low	Limited	Known
				Not known
3 (insufficient)	3 (insufficient)	None	None	Not known

Household Water Chlorination

TANG Framework	
What works is known or what works is not known	Known: at least 4 RCTs plus other controlled trials, pooled analyses show reduction in diarrhea of RR = 0.61 or 0.71
Association (defined in TANG as RR ≥ 2)	If diarrhea is the outcome, then it is low (RR = 0.61 – 0.71). If fecal coliforms is the outcome, then it is high (RR = 0.20).
Repeatability (and/or consistency*)	Efficacy/effectiveness shown in 4 continents. Seems like this should be "repeatability has been widely demonstrated".
How it works	Known . Chlorine kills germs, germs cause diarrhea.
Overall classification	Grade 1 (strong) level 1[if fecal coliforms is the outcome], Grade 2B (possible) level 1 if diarrhea is the outcome

Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared to usual care

TANG Framework	
What works is known or what works is not known	Known: 7 RCTs, 6 of which favor the intervention, 2 of which produced statistically significant results, the pooled effect is statistically significant.
Association (defined in TANG as RR ≥ 2)	Low (RR= 0.86)
Repeatability (and/or consistency*)	Effectiveness has been studied in 6 different countries, but all but one is in southeast Asia.
How it works	Not known. Interventions have been heterogeneous, there are no data about direct process-outcome links.
Overall classification	Grade 2b Level 2 (if repeatability outside southeast Asia is acceptable) Grade 2c Level 2 (if repeatability outside southeast Asia is considered acceptable)

PMTCT

TANG Framework	
What works is known or what works is not known	Known: Several high quality meta-analyses show significant results.
Association (defined in TANG as RR ≥ 2)	Yes, High; Transmission reduced by 40 to 80% depending on regimen.
Repeatability (and/or consistency*)	Wide. Consistent results. Has been studied in many RCTs in Sub-Saharan Africa, Southeast Asia, and Latin America, after first being tested in developed countries.
How it works	Mechanism is well understood at the molecular basis.
Overall classification	Grade 1 (Strong) Level 1

* Consistency is used as the term in the abstract, but repeatability is the term in the text, and these seem like potentially different constructs.

GRADE Framework[2]

Table B.2 A summary of GRADE's approach to rating quality of evidence

Study Design	Initial quality of a body of evidence	Lower if	Higher if	Quality of a body of evidence
Randomized Trials	High ⇒	Risk of Bias - 1 Serious - 2 Very serious Inconsistency - 1 Serious - 2 Very serious	Large effect + 1 Large + 2 Very large Dose response + 1 Evidence of a gradient	High (four plus: ⊕ ⊕ ⊕ ⊕) Moderate (three plus: ⊕ ⊕ ⊕◯)
Observational studies	Low ⇒	Indirectness - 1 Serious - 2 Very serious Imprecision - 1 Serious - 2 Very serious Publication bias - 1 Likely - 2 Very likely	All plausible residual confounding + 1 Would reduce a Demonstrated effect + 1 Would suggest a Spurious effect if no effect was observed	Low (two plus: ⊕ ⊕ ◯◯) Very low (one plus: ⊕◯◯◯)

Table B.3 Significance of the four levels of evidence

Quality Level	Current Definition
High	We are very confident that the true effect lies close to that of the estimate of the effect
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very Low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Household Water Chlorination

No. of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations
Water Quality						
5 +	Controlled Trials	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Strong association
Diarrhea						
5+	Controlled Trials	Serious or very serious limitations (blinding)	No serious inconsistency	No serious indirectness	No serious imprecision	None
Summary of findings						
	Effect		**Quality**		**Importance**	
Water Quality						
	RR = 0.20 (95% CI 0.13 – 0.30)		High ⊕⊕⊕⊕		Important	
Diarrhea						
	RR = 0.71 (95% CI 0.58 – 0.87)		Moderate : ⊕⊕◐		Critical	

Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared to usual care

No. of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations
Reduce Morbidity / Mortality						
7	Randomized Trials	Serious risk of bias	Serious inconsistency	No serious indirectness	No serious imprecision	None
Summary of findings						
	Effect		**Quality**		**Importance**	
Reduce Morbidity / Mortality						
	RR = 0.86 (95% CI 0.75 – 0.99)		Low ⊕⊕○○		Critical	

PMTCT

No. of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations
HIV Infection in Child						
15	All RCTS	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	None
Summary of findings						
	Effect		**Quality**		**Importance**	
	Transmission of HIV reduced by 40% to 80% depending on regimen		High ⊕⊕⊕⊕		High. Huge population affected in Sub Saharan Africa.	

HASTE Framework[3]

Table B.4 Highest Attainable Standard of Evidence System for HIV Interventions (HASTE)

Grade Level	Strength of Recommendation	Explanation
Grade 1	Strong	• Efficacy is Consistent • Biologically plausible • Implementation Data Available
Grade 2a	Probable	• Limited Efficacy Data • Biologically Plausible
Grade 2b	Possible	• Efficacy is inconsistent • Biologically plausible • Consensus from Implementation Science Data
Grade 2c	Pending	• Ongoing Definitive Trials • Plausibility
Grade 3	Insufficient	• Inconsistent Data • Undefined Plausibility • Paucity of Implementation Science Data
Grade 4	Inappropriate	• Consistent data demonstrating lack of efficacy • Consensus from Implementation Data of Inappropriate Intervention

Household Water Chlorination

HASTE Framework	
Efficacy is Consistent	Yes, based on the forest plot in 2 meta-analyses
Implementation science data	We can only use what is in the included studies or implementation science studies from the Arnold list of excludes. We rate this as "available", certainly implementation is described better in these studies than in many studies medical interventions.
Biologically Plausible	Yes
Overall classification based on HASTE	Grade 1 Strong

Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared to usual care

HASTE Framework	
Efficacy is Consistent	Yes, based on the forest plot in the Cochrane Review.
Implementation science data	We can only use what is in the included studies in the Cochrane Review. These studies give quite detailed explanations of how the interventions were implemented. We rate this as "available".
Biologically Plausible	Unclear
Overall classification based on HASTE	Unclear how to clarify this. "Undefined plausibility" appear in Grade 3, insufficient, but the other descriptor in that case say "inconsistent data" and "paucity of implementation science data") don't accurately characterize the evidence.

PMTCT

HASTE Framework	
Efficacy is consistent	Yes, based on forest plots in multiple meta-analyses.
Implementation science data	Exists, focuses on barriers, costs
Biologically Plausible	Yes
Overall classification based on HASTE	Grade 1, Strong

Community Preventive Services Task Force Framework[4]

Table B.5 Assessing the strength of a body of evidence on effectiveness of population-based interventions in the *Guide to Community Preventive Services*

Evidence of effectiveness[a]	Execution—good or fair[b]	Design Suitability—Greatest, moderate, or least	Number of studies	Consistent[c]	Effect size[d]	Expert opinion[e]
Strong	Good	Greatest	At Least 2	Yes	Sufficient	Not used
	Good	Greatest or Moderate	At Least 5	Yes	Sufficient	Not used
	Good or Fair	Greatest	At Least 5	Yes	Sufficient	Not used
	Meet Design, Execution, Number and Consistency Criteria for Sufficient But Not Strong Evidence				Large	Not used
Sufficient	Good	Greatest	1	Not Applicable	Sufficient	Not used
	Good or Fair	Greatest or Moderate	At Least 3	Yes	Sufficient	Not used
	Good or Fair	Greatest, Moderate or Least	At Least 5	Yes	Sufficient	Not used
Expert Opinion	Varies	Varies	Varies	Varies	Sufficient	Supports a Recommendation
Insufficient[f]	A. Insufficient	Designs or Execution	B. Too Few Studies	C. Inconsistent	D. Small	E. Not Used

[a]The categories are not mutually exclusive; a body of evidence meeting criteria for more than one of these should be categorized in the highest possible category.
[b]Studies with limited execution are not used to assess effectiveness.
[c]Generally consistent in direction and size.
[d]Sufficient and large effect sizes are defined on a case-by-case basis and are based on Task Force opinion.
[e]Expert opinion will not be routinely used in the *Guide* but can affect the classification of a body of evidence as shown.
[f]Reasons for determination that evidence is insufficient will be described as follows: A. Insufficient designs or executions, B. Too few studies, C.Inconsistent. D. Effect size too small, E. Expert opinion not used. These categories are not mutually exclusive and one or more of these will occur when a body of evidence fails to meet the criteria for strong or sufficient evidence.

Table B.6 Suitability of study design for assessing effectiveness in the Guide to *Community Preventive Services*

Suitability	Attributes
Greatest	Concurrent comparison groups and prospective measurement of exposure and outcome
Moderate	All retrospective designs or multiple pre or post measurements but no concurrent comparison group
Least	Single pre and post measurements and no concurrent comparison group or exposure and outcome measured in a single group at the same point in time

Household Water Chlorination

Community Preventive Services Task Force Framework	
Household Water Chlorination	
Water Quality	
Execution	Good or Fair
Design suitability	Greatest
Number of studies	At least 5
Consistent	Generally consistent in direction and size
Effect size	Sufficient
Overall evidence of effectiveness	Strong
Diarrhea	
Execution	Good or Fair
Design suitability	Greatest
Number of studies	At least 5
Consistent	Generally consistent in direction and size
Effect size	Sufficient
Overall evidence of effectiveness	Strong

Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared to usual care

Community Preventive Services Task Force Framework	
Execution	Good or Fair
Design suitability	Greatest
Number of studies	At least 5
Consistent	Generally consistent in direction and size
Effect size	Sufficient
Overall evidence of effectiveness	Strong

PMTCT

Community Preventive Services Task Force Framework	
Execution	Good
Design suitability	Greatest (RCTs)
Number of studies	At least 5
Consistent	Yes
Effect size	Large
Overall evidence of effectiveness	Strong

Australian NHMRC Framework[5]

Table B.7 Body of evidence matrix

Component	A	B	C	D
	Excellent	**Good**	**Satisfactory**	**Poor**
Evidence base[a]	one or more level I studies with a low risk of bias or several level II studies with a low risk of bias	one or two level II studies with a low risk of bias or a SR/several level III studies with a low risk of bias	one or two level III studies with a low risk of bias, or level I or II studies with a moderate risk of bias	level IV studies, or level I to III studies/SRs with a high risk of bias
Consistency[b]	all studies consistent	most studies consistent and inconsistency may be explained	some inconsistency reflecting genuine uncertainty around clinical question	evidence is inconsistent
Clinical impact	very large	substantial	moderate	slight or restricted
Generalizability	population/s studied in body of evidence are the same as the target population for the guideline	population/s studied in the body of evidence are similar to the target population for the guideline	population/s studied in body of evidence differ to target population for guideline but it is clinically sensible to apply this evidence to target population[c]	population/s studied in body of evidence differ to target population and hard to judge whether it is sensible to generalise to target population
Applicability	directly applicable to Australian healthcare context	applicable to Australian healthcare context with few caveats	probably applicable to Australian healthcare context with some caveats	not applicable to Australian healthcare context

SR = systematic review; several = more than two studies

[a] Level of evidence determined from the NHMRC evidence hierarchy

[b] If there is only one study, rank this component as 'not applicable'.

[c] For example, results in adults that are clinically sensible to apply to children OR psychosocial outcomes for one cancer that may be applicable to patients with another cancer

Table B.8 Definition of NHMRC grades of recommendations

Grade of recommendation	Description
A	Body of evidence can be trusted to guide practice
B	Body of evidence can be trusted to guide practice in most situations
C	Body of evidence provides some support for recommendation(s) but care should be taken in its application
D	Body of evidence is weak and recommendation must be applied with caution

Table B.9 NHMRC Evidence Hierarchy: designations of 'levels of evidence' according to type of research question (including explanatory notes)

Level	Intervention[a]	Diagnostic Accuracy[b]	Prognosis	Aetiology[c]	Screening Intervention
I[d]	A systematic review of level II studies	A systematic review of level II studies	A systematic review of level II studies	A systematic review of level II studies	A systematic review of level II studies
II	A randomized controlled trial	A study of test accuracy with: an independent, blinded comparison with a valid reference standard,[e] among non-consecutive persons with a defined clinical presentation[f]	A prospective cohort study[g]	A prospective cohort study	A randomized controlled trial
III-1	A pseudorandomized controlled trial (i.e. alternate allocation or some other method)	A study of test accuracy with: an independent, blinded comparison with a valid reference standard,[e] among non-consecutive persons with a defined clinical presentation[f]	All or none[h]	All or none[h]	A pseudorandomized controlled trial (i.e. alternate allocation or some other method)
III-2	A comparative study with concurrent controls: ▪ Non-randomized, experimental trial[i] ▪ Cohort study ▪ Case-control study ▪ Interrupted time series with a control group	A comparison with reference standard that does not meet the criteria required for Level II and III-1 evidence	Analysis of prognostic factors amongst persons in a single arm of a randomized controlled trial	A retrospective cohort study	A comparative study with concurrent controls: ▪ Non-randomized, experimental trial ▪ Cohort study ▪ Case-control study
III-3	A comparative study without concurrent controls: ▪ Historical control study ▪ Two or more single arm Study[j] ▪ Interrupted time series without a parallel control group	Diagnostic case-control Study[f]	A retrospective cohort study	A case-control study	A comparative study without concurrent controls: ▪ Historical control study ▪ Two or more single arm Study
IV	Case series with either post-test or pre-test/post-test outcomes	Study of diagnostic yield (no reference standard)[k]	Case series, or cohort study of persons at different stages of disease	A cross-sectional study or case series	Case series

Explanatory notes

a. Definitions of these study designs are provided on pages 7-8 How to use the evidence: assessment and application of scientific evidence
(NHMRC 2000b) and in the accompanying Glossary.

b. These levels of evidence apply only to studies of assessing the accuracy of diagnostic or screening tests. To assess the overall effectiveness of a diagnostic test there also needs to be a consideration of the impact of the test on patient management and health outcomes (Medical Services Advisory Committee 2005, Sackett and Haynes 2002). The evidence hierarchy given in the 'Intervention' column should be used to assess the impact of a diagnostic test on health outcomes relative to an existing method of diagnosis/comparator test(s). The evidence hierarchy given in the 'Screening' column should be used to assess the impact of a screening test on health outcomes relative to no screening or opportunistic screening.

c. If it is possible and/or ethical to determine a causal relationship using experimental evidence, then the 'Intervention' hierarchy of evidence should be utilised. If it is only possible and/or ethical to determine a causal relationship using observational evidence (eg. cannot allocate groups to a potential harmful exposure, such as nuclear radiation), then the 'Aetiology' hierarchy of evidence should be utilised.

d. A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of level II evidence. Systematic reviews of level II evidence provide more data than the individual studies and any meta-analyses will increase the precision of the overall results, reducing the likelihood that the results are affected by chance. Systematic reviews of lower level evidence present results of likely poor internal validity and thus are rated on the likelihood that the results have been affected by bias, rather than whether the systematic review itself is of good quality. Systematic review quality should be assessed separately. A systematic review should consist of at least two studies. In systematic reviews that include different study designs, the overall level of evidence should relate to each individual outcome/result, as different studies (and study designs) might contribute to each different outcome.

e. The validity of the reference standard should be determined in the context of the disease under review. Criteria for determining the validity of the reference standard should be pre-specified. This can include the choice of the reference standard(s) and its timing in relation to the index test. The validity of the reference standard can be determined through quality appraisal of the study (Whiting et al 2003).

f. Well-designed population based case-control studies (eg. population based screening studies where test accuracy is assessed on all cases, with a random sample of controls) do capture a population with a representative spectrum of disease and thus fulfil the requirements for a valid assembly of patients. However, in some cases the population assembled is not representative of the use of the test in practice. In diagnostic case-control studies a selected sample of patients already known to have the disease are compared with a separate group of normal/healthy people known to be free of the disease. In this situation patients with borderline or mild expressions of the disease, and conditions mimicking the disease are excluded, which can lead to exaggeration of both sensitivity and specificity. This is called spectrum bias or spectrum effect because the spectrum of study participants will not be representative of patients seen in practice (Mulherin and Miller 2002).

g. At study inception the cohort is either non-diseased or all at the same stage of the disease. A randomised controlled trial with persons either non-diseased or at the same stage of the disease in both arms of the trial would also meet the criterion for this level of evidence.

h. All or none of the people with the risk factor(s) experience the outcome; and the data arises from an unselected or representative case series which provides an unbiased representation of the prognostic effect. For example, no smallpox develops in the absence of the specific virus; and clear proof of the causal link has come from the disappearance of small pox after large-scale vaccination.

i. This also includes controlled before-and-after (pre-test/post-test) studies, as well as adjusted indirect comparisons (ie. utilise A vs B and B vs C, to determine A vs C with statistical adjustment for B).

j. Comparing single arm studies ie. case series from two studies. This would also include unadjusted indirect comparisons (ie. utilise A vs B and B vs C, to determine A vs C but where there is no statistical adjustment for B).

k. Studies of diagnostic yield provide the yield of diagnosed patients, as determined by an index test, without confirmation of the accuracy of this diagnosis by a reference standard. These may be the only alternative when there is no reliable reference standard.

Note A: Assessment of comparative harms/safety should occur according to the hierarchy presented for each of the research questions, with the proviso that this assessment occurs within the context of the topic being

assessed. Some harms (and other outcomes) are rare and cannot feasibly be captured within randomised controlled trials, in which case lower levels of evidence may be the only type of evidence that is practically achievable; physical harms and psychological harms may need to be addressed by different study designs; harms from diagnostic testing include the likelihood of false positive and false negative results; harms from screening include the likelihood of false alarm and false reassurance results.

Note B: When a level of evidence is attributed in the text of a document, it should also be framed according to its corresponding research question eg. level II intervention evidence; level IV diagnostic evidence; level III-2 prognostic evidence.

Note C: Each individual study that is attributed a "level of evidence" should be rigorously appraised using validated or commonly used checklists or appraisal tools to ensure that factors other than study design have not affected the validity of the results.

Source: Hierarchies adapted and modified from: NHMRC 1999; Bandolier 1999; Lijmer et al. 1999; Phillips et al. 2001.

Household Water Chlorination

Australian NHMRC Framework	Water Quality	Diarrhea
Evidence of effectiveness	A - one or two level 1 (actually two) studies with a low risk of bias	A
Consistency	A - All studies consistent	B - most studies consistent and inconsistency may be explained
Clinical Impact	Substantial	Substantial
Generalizability	B	C
Applicability	C - For this application we'll say "probably applicable with some caveats"	C - For the purpose we'll assume a rural area in a MIC
Overall classification based on Australian NHMRC	B	C

Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared to usual care

Australian NHMRC Framework	
Evidence of effectiveness	B – One or two (actually 1) level 2 studies with a low risk of bias
Consistency	B – Most studies consistent and inconsistency may be explained
Clinical Impact	C - Moderate
Generalizability	B - population/s studied are similar to the target population
Applicability	C – Probably applicable with some caveats
Overall classification based on Australian NHMRC	B – "Good"

PMTCT

Australian NHMRC Framework	
Evidence of effectiveness	A - Excellent, several high quality systematic reviews of RCTs
Consistency	B – Good. Very consistent – only inconsistencies in effect are explained. Studies with lower effect sizes conducted in breastfeeding women
Clinical Impact	A – Very large, possibly millions worldwide
Generalizability	B – Has been studied in Latin America, Asia, Sub-Saharan Africa
Applicability	C – In clinical trials, drugs are provided for free. In reality, free meds aren't available to everyone who needs them.
Overall classification based on	B - Body of evidence can be trusted to guide practice in most

Australian NHMRC Framework	
Australian NHMRC	situations
Limitations	Doesn't consider cost directly, unless taken into consideration as part of "applicability"

NHS Health Development Agency Framework[6]

Table B.10 Pilot public health evidence grading scheme: classification of recommendation

Class	Basis for decision*
A [PH]	At least one 1++ study or consistent findings in a body of studies** principally rated as 1+ for efficacy***, with strong or moderate evidence of corroboration OR Consistent findings in a body of 2++ studies for efficacy, with strong evidence of corroboration
B [PH]	At least one 1++ study or consistent findings in a body of studies principally rated as 1+ for efficacy, with limited/ no evidence of corroboration OR A single 1+ study for efficacy, with strong or moderate evidence of corroboration OR A single 2++ study or consistent findings in a body of studies principally rated as 2+ for efficacy, with strong evidence of corroboration OR Consistent findings in a body of studies principally rated as 2++ for efficacy, with moderate evidence of corroboration
C [PH]	Consistent findings in a body of studies principally rated as 2++ for efficacy, with limited/no evidence of corroboration OR A single 2++ study or consistent findings in a body of studies principally rated 2+ for efficacy, with moderate evidence of corroboration OR A single 2+ study for efficacy, with strong evidence of corroboration OR A body of level 3 or 4 evidence for efficacy, with strong evidence of corroboration
D [PH]	A single 2++ study or consistent findings in a body of studies principally rated 2+ for efficacy, with limited/no evidence of corroboration OR A single 2+ study for efficacy, with moderate evidence for corroboration OR A body of level 3 or 4 evidence of efficacy, with moderate/limited evidence of corroboration OR Formal consensus
D [GPP]	A recommendation based on experience of best practice by health professionals and expert groups

*See Tables 2 and 3 for key to study type, quality and strength of evidence.
**Body of studies = 3 or more, or a systematic review.
***For national environmental/socio-political interventions, a body of 2+ studies is acceptable.
[PH] public health; [GPP] Good Practice Point.
Source: adapted from SIGN (2001).

Table B. 11 Evidence of the efficacy of an intervention – did it work?

Level of evidence	Type of evidence
1++	High quality meta-analyses, systematic reviews of RCTs (including cluster RCTs), or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias
1–*	Meta-analyses, systematic reviews of RCTs, or RCTs with a high risk of bias
2++	High quality systematic reviews of, or individual high quality non-randomised intervention studies (controlled non-randomised trial, controlled before-and-after, interrupted time series), comparative cohort and correlation studies with a very low risk of confounding, bias or chance
2+	Well conducted, non-randomised intervention studies (controlled non-randomised trial, controlled before-and-after, interrupted time series), comparative cohort and correlation studies with a low risk of confounding, bias or chance
2–*	Non-randomised intervention studies (controlled non-randomised trial, controlled before-and-after, interrupted time series), comparative cohort and correlation studies with a high risk of confounding, bias or chance
3	Non-analytical studies (eg case reports, case series)
4	Expert opinion, formal consensus

*Studies with a level of evidence (–) should not be used as basis for making recommendations.
Source: adapted from SIGN (2001).

Table B. 12 Evidence for corroboration – will it work? (evidence to support implementation in the UK today) and does it matter? (evidence of salience and relevant outcomes for UK populations today)

Strength of evidence	Type of evidence
Strong	Consistent findings in two or more studies of ++ quality carried out within the UK and applicable* to the target population, providing evidence on salience and implementation
Moderate	One ++ study or consistent findings in two or more studies of + quality carried out within the UK and applicable to the target population OR
	Two or more ++ studies from outside the UK but applicable to the target population, providing evidence on salience and implementation
Limited	Only one + study from the UK, two or more studies with inconsistent findings (on balance providing evidence of benefit or harm) or studies of + quality from outside the UK
No evidence	No study of acceptable quality, inconsistent findings (on balance providing no useful evidence) or no relevant research available

Household Water Chlorination

NHS Health Development Agency Framework	
Evidence of efficacy – did it work?	Class 1++ or class 1+ for water quality Class 2++ for diarrhea
Evidence for corroboration – will it work? (evidence to support implementation) And does it matter? (evidence of salience to the target population)	If the studies are ++ (meaning "high quality" with a very low risk of bias, confounding or chance); if not this strength then "limited" if the studies are only "well conducted" with a low risk of bias, confounding, or chance.
Overall classification using NHS Health Development Agency	"A" for water quality "B" for diarrhea

Lay health workers in primary or community health care to reduce mortality and morbidity in children under age 5, compared to usual care

NHS Health Development Agency Framework	
Evidence of efficacy – did it work?	A single 2++ study (Cochrane meta-analysis) with moderate evidence of corroboration.
Evidence for corroboration – will it work? (evidence to support implementation) And does it matter? (evidence of salience to the target population)	A single 2++ study (Cochrane meta-analysis) with moderate evidence of corroboration.
Overall classification using NHS Health Development Agency	"C"

PMTCT

NHS Health Development Agency Framework	
Evidence of efficacy – did it work?	1++, several high quality meta-analyses of RCTs
Evidence for corroboration – will it work? (evidence to support implementation) And does it matter? (evidence of salience to the target population)	Studies are conducted in the target population in low income countries.
Overall classification using NHS Health Development Agency	"A"

References

1. Tang KC, Choi BC, Beaglehole R. Grading of evidence of the effectiveness of health promotion interventions. J Epidemiol Community Health. 2008;62(9):832-4. PMID: 18701736.
2. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401-6. PMID: 21208779.
3. Baral S, Wirtz A, Sifakis F, et al. "The Highest Attainable Standard of Evidence (HASTE) for HIV/AIDS Interventions: Towards A Public Health Approach to Defining Evidence. Public Health Reports, in press.
4. Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based Guide to Community Preventive Services--methods. The Task Force on Community Preventive Services. Am J Prev Med. 2000;18(1 Suppl):35-43. PMID: 10806978.
5. National Health and Medical Research Council. NHMRC levels of evidence and grades for recommendations for developers of guidelines. Australia; December 2009.
6. Wightman A, Ellis S, Cullum A, et al. Grading evidence and recommendations for public health interventions: developing and piloting a framework. Health Development Agency. 2005 http://www.nice.org.uk/nicemedia/docs/grading_evidence.pdf Accessed Dec 2012.

Appendix C. Pilot Test of Implementation Reporting Criteria on Three Published Studies for each of Three Exemplars

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	We postulated that an intervention based on a socioculturally contextualised approach of behavior change management systematically applied to modifiable, high-risk newborn-care practices, with an emphasis on hypothermia, within a community with a high neonatal mortality rate could lead to improved care practices and reduced mortality. (*This seems to indicate the intervention was developed by the research teams – but later, there is this text:* Qualitative research activities provided the evidence base for investigators and community members to codevelop the intervention strategy, which underwent further refinement based on findings of trials of improved practices. *so should this be considered internally developed?)*	Fair
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also	In a study in Maharashtra, India, Bang and colleagues reported a 62–70% reduction in the neonatal mortality rate, and attributed 93% of the reduction to active management of sick newborn babies and 7% to primary prevention. Baqui and colleagues reported that an adaptation of this approach in Bangladesh in an effectiveness trial had half the effect (34% reduction) on neonatal mortality. Manandhar and co-workers3 tested a different approach in Nepal with a community-based participatory action-cycle with no prespecified intervention package, in which women's groups identified priorities and implemented local solutions, and reported improvements in care practices, care-seeking, and a 30% reduction in neonatal mortality rate. We postulated that an intervention based on a socioculturally contextualised approach of behavior change management systematically applied to modifiable, high-risk newborn-care practices, with an emphasis on hypothermia, within a community with a high neonatal mortality rate could lead to improved care practices and reduced mortality. *(Identifies both a framework for why it should work and prior data)*	Good

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."		
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	Our preliminary qualitative field work showed that individual behaviours were influenced by collective behaviours and social norms, and sustained by a complex, multilevel network of relationships within the community. We therefore developed a multilevel strategy targeting: community stakeholders, newborn stake holders, and households with immediate support groups. At each level, the target group consisted of individuals who were identified to have key roles as influencers, decision makers, supporters, and practitioners of newborn care and normative behaviour within the community. The support of community stake holders such as village heads, community leaders, respected members, priests, and teachers was crucial in building trust with the community and ensuring acceptance of the programme. The newborn stakeholder target group included traditional newborn-care providers and birth attendants, unqualified medical practitioners, and, to a lesser extent, health system workers, some of whom had strategic access to the newborn and mother during post-partum confinement, were perceived by the community as domain experts, and played an active part in sustaining targeted practices. Health system workers such as auxillary nurse midwives were engaged only at the community level as part of newborn stakeholder group meetings in order to keep contamination of the intervention into control clusters to a minimum. The household target group included the pregnant woman or mother, who was the primary care provider, but usually not empowered to make decisions; the mother-in-law, who was usually the key decision maker on newborn-care practices; other female members who played supportive roles; and male members, including the father-in-law and husband, who controlled access to the household, made financial and logistical arrangements, and influenced care-seeking decisions. The family's immediate support group included neighbours and relatives who influenced family behaviours and helped with deliveries. *(Indicates extensive field work to fit intervention components to the needs of the population)*	Good
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:**	Not explicitly mentioned	None

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.		
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	The primary enablers of behaviour change were paid (US$35–40 per month) community-based health workers, the Saksham Sahayak (n=26), who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders.	Good
b. Characteristics of the recipients	At each level, the target group consisted of individuals who were identified to have key roles as influencers, decision makers, supporters, and practitioners of newborn care and normative behaviour within the community.	Good
c. The setting	The state of Uttar Pradesh, India, accounts for a quarter of India's neonatal deaths and for 8% of those worldwide, and shares similar sociocultural, demographic, and health system characteristics with other high-mortality Indian states and south Asian countries. The study was done in Shivgarh, a rural block in Uttar Pradesh, with a population of 104 123 divided into 39 village administrative units . Socioeconomic indicators are among the lowest in the state. The formal health-care system in Shivgarh consists of a community health centre and two primary health centres operated by trained physicians and paramedical staff supported by 18 auxiliary nurse midwives, who are outreach workers catering to a population of 6000–7000 each, and trained to deliver babies, and provide vaccinations and antenatal check-ups. Care-seeking from them, however, is low.	Good
d. The mode of delivery (such as face-to-face)	The intervention was delivered from January, 2004, to May, 2005. Saksham Sahayaks first engaged with community stakeholders in community meetings	Good

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
	to seek their approval, sensitise them towards the importance of their role in newborn survival, encourage shared learning, and create a supportive environment (figure 1, webtable 2). Folk song group meetings, where messages to promote behaviour change were incorporated into folk songs, were held by Saksham Sahayaks on a monthly basis with participants from diverse target groups. They also held separate monthly meetings with newborn-care stakeholders and with community volunteers to discuss experiences, challenges, and strategies.	
e. The intensity of the intervention/program (such as the contact time with participants)	This process was accomplished through 3-monthly cycles of door-to-door household visits by Saksham Sahayaks, self-reporting by pregnant women, and information provided by community volunteers. An antenatal visit was planned for 60 days before the expected date of delivery and another for 30 days before the expected date of delivery to provide ample time for effective behaviour change negotiation, ensure birth preparedness, and build trust with the family to negotiate subsequent entry into the room of confinement after delivery for postnatal visits (webtable 3). Post-partum confinement was a universal practice, and coincided with the initiation of almost all the targeted practices and occurrence of most newborn deaths. As some of the new practices were skill-based,	Good
f. The duration (such as the number of sessions and their spacing interval over a given period)	the first postnatal visit was planned within 24 h of the delivery and the second postnatal visit was planned on day 3 (webtable 3). In case of sick neonates, no treatment was provided, but families were advised to seek care at the nearest health facility.	Good
g. Adherence or fidelity to delivery protocols	Coverage of household visits by *Saksham Sahayaks* was calculated as the ratio of total visitations recorded during the study period to the total number of women eligible for the visitations. For coverage on antenatal visits, all pregnancies were considered eligible and for coverage on postnatal visits, all women with at least one liveborn baby were considered eligible for the visits. Household visits by newborn-care stakeholders and community volunteers in the absence of *Saksham Sahayaks* were not recorded. The monthly coverage of group meetings was based on monitoring reports by *Saksham Sahayaks*.	Good
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3]	The primary enablers of behaviour change were paid (US$35–40 per month) community-based health workers. *(Only cost data identified)*	Poor / None

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
Explanation/Example: The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.		
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	Participatory social mapping of all villages in the study area provided an introduction to the community, initiated the process of collaborative engagement, served to identify community resources for newborn health, and facilitated the planning of home visitations and group interventions. Qualitative research activities provided the evidence base for investigators and community members to codevelop the intervention strategy, which underwent further refinement based on findings of trials of improved practices. Domiciliary care practices were mapped against the existing evidence base of risk factors for neonatal mortality and morbidity. Practices that were assessed to be potentially harmful, preventable, within community control, and amenable to change were selected for behavioural modification (webtable 1). *(Field work to identify potential barriers and facilitators of implantation)*	Good
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor	No text was found about facilitators and barriers to the actual implementation.	Poor / None

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."		
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	Available in a web appendix.	Good
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's	No text was found describing the assessment of implementation.	Poor / None

Lay Health Workers

Table C.1.1 - Criterion Table for Kumar, Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial, Lancet, 2008; 372: 1151-62[1]

Criterion	Example of text related to this criterion	Rating
implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."		

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	This project is an EU INCO-DEV funded collaboration between the Heidelberg University (Germany), Karolinska Institute (Sweden), Muhimbili University College of Health Sciences (Tanzania) and Centre de Recherche en Santé de Nouna (Burkina Faso) called MAMOP project (Improving the management of childhood Malaria: an experiment to bridge the gap between Mothers and health care Providers). *(Indicates the intervention was developed externally)*	Good
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	No text was found.	Poor / None

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	No text was found.	Poor / None
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	Not explicitly stated.	Poor / None

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	Inclusion criteria for group leaders used by the communities were permanent residency in the sub-village, age 30–50 years, honesty, and respect by the community. A five days training course for the health workers of participating health centres was conducted by one of the investigators (FS).	Fair
b. Characteristics of the recipients	The intervention was targeted at three groups: health workers (nurses) from five peripheral health centres (Toni, Dara, Bourasso, Lekuy, Koro), women group leaders, and the main care takers (usually the mothers) of preschool children.	Fair
c. The setting	The study was implemented in the rural part of the research zone of the Centre de Recherche en Santé de Nouna (CRSN) in Nouna Health District, north-western Burkina Faso (Figure 1). The Nouna area is a dry orchard savannah, populated mainly by subsistence farmers of different ethnic groups. Malaria is holoendemic but highly seasonal, and the transmission intensity varies between 100 and 1000 infective bites per person and year between study villages. Formal health services in the study area are provided by a limited number of rural health centres and the district hospital in Nouna town. Villagebased health centres are usually equipped with two nurses and one mid-wife and do outreach work in the surrounding surrounding 7–10 villages under their responsibility. Malaria control is mainly based on home treatment with CQ, which has been shown to be still sufficiently effective in 2001, and on malaria prophylaxis for pregnant women. Untreated mosquito nets have been used in the area for a long time, but insecticide-treated nets (ITN) were only recently introduced in the frame of an effectiveness study. Communities in the study area have been shown to be quite well organized with regard to risk sharing mechanisms. In particular women groups with a focus on mutual agricultural support traditionally exist in all villages.	Good

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
d. The mode of delivery (such as face-to-face)	Although not stated, it is implied that the intervention is developed face-to-face.	Fair
e. The intensity of the intervention/program (such as the contact time with participants)	No text was found.	Poor / None
f. The duration (such as the number of sessions and their spacing interval over a given period)	No text was found.	Poor / None
g. Adherence or fidelity to delivery protocols	The only data on adherence or fidelity to delivery protocols was self-reported treatment with chloroquine listed in Table 4. Some components are described in great detail. Reproductions of charts and pictures provided to recipients are presented, however other components such as the training given to women group leaders were only briefly described or not at all.	Fair
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	The only cost data reported was the cost of chloroquine.	Poor / None

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	The only description of population needs was that malaria is holoendemic to the area.	Poor / None
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	No text found.	Poor / None

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	Some components are described in great detail. Reproductions of charts and pictures provided to recipients are presented, however other components such as the training given to women group leaders were only briefly described or not at all.	Fair

Lay Health Workers

Table C.1.2 - Criterion Table for Koyate B.; Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial, Malaria Journal 2008; 7(50)[5]

Criterion	Example of text related to this criterion	Rating
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	No text found.	Poor / None

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	Together with the Bangladesh Rural Advancement Committee (BRAC), Mitra and Associates, Ecuadorian and Bangladeshi physicians, nurse-midwives, and KMC experts, the study team adapted KMC so that it can be feasibly implemented as a community-based intervention (CKMC). (*This seems to imply that the intervention was externally developed)*	Good

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	Kangaroo mother care (KMC) is a method whereby the hospital-born stabilized LBW newborn is placed in skin-to-skin (STS) contact on the mother's breast to promote thermal regulation, breastfeeding, and maternal–infant bonding. Traditional KMC reduces the incidence of morbidity but not mortality in LBW infants, because it is generally applied to clinically stabilized newborns and most neonatal mortality occurs in the first 2 days of life before stabilization. A single adequately designed study found a 43% (not statistically significant) lower infant mortality rate (IMR) associated with traditional KMC. Two small African studies of early (as soon as possible after birth) KMC in hospitals with little neonatal intensive care capacity reported reduced mortality within 24 hours of birth and before discharge; however, important differences in study group characteristics were not controlled in analysis. *(Describes the rational for why the intervention should work (Thermoregulation, breast feeding, bonding) and describes prior data from similar studies)*	Poor / None
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	No text was found.	Poor / None

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	Not explicitly stated.	Poor / None

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	A physician who had participated in the pilot study trained 12 BRAC supervisors and, along with 1 supervisor, trained all 63 community nutrition workers and their 25 NNP supervisors serving the intervention group in 5 groups of 6 to 22 people during a 2-month period.	Fair
b. Characteristics of the recipients	A detailed table of characteristics of the recipients is included in Table 1.	Good
c. The setting	Bangladesh has a population of 140 000 000 administratively divided into 6 divisions that, combined, contain 64 districts and 496 subdistricts called upazilas, each of which has a capitol city. Each study subdistrict contains 8 to 14 unions, and each union contains 5 to 25 villages. The sample includes the 42 unions that participated in the NNP, all that are supervised by our study partner BRAC in the Dhaka and Sylhet divisions. Dhaka and Sylhet divisions are located in northern Bangladesh, where NMR was 5.2% and 8.2% and IMR was 11.5% and 16.2%, the highest in the nation, when the study was designed.	Good
d. The mode of delivery (such as face-to-face)	Although not stated, it is implied that the intervention is developed face-to-face.	Fair
e. The intensity of the intervention/program (such as the contact time with participants)	No text was found.	Poor / None
f. The duration (such as the number of sessions and their spacing interval over a given period)	No text was found.	Poor / None

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
g. Adherence or fidelity to delivery protocols	The training and intervention delivery processes that were used in the pilot study were only partially transferred to the full trial, with unplanned substitution of experienced trainers with individuals who were not trainers, less frequent contact between community workers and mothers in the last month of pregnancy, and unplanned emphasis on CKMC for LBW infants. The nested qualitative study found that >35% of CKMC women were erroneously taught that STS was to be provided to LBW or preterm infants rather than to all infants, and only 30% were correctly taught to hold all infants STS. Less than 40% of CKMC mothers were taught to provide CKMC to infants who were ill, and >25% were erroneously taught to breastfeed on schedule (not on demand). Women had numerous views about the number of hours and days they should provide STS, indicating that they received variable and frequently incorrect messages from the community workers and supervisors.	Good
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	All participating community workers received $7.50 a month. *(Only cost data identified)*	Poor / None

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	No text was found.	Poor / None
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	Field visits confirmed that some of those who were employed to conduct the CKMC training believed that CKMC was intended for small infants. Thus, some intervention group mothers may not have provided CKMC because they were mistakenly taught that CKMC is for small infants. CKMC implementation and effect depend on both the quality of CKMC training and the mother's behavior modification, making it difficult to know whether the intervention does not have effect in larger, more mature infants or whether the uptake was suboptimal as a result of insufficient training or poor maternal adherence. *(This indicates that insufficiently trained CKMC trainers decreased the effectiveness of the intervention)*	Good

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	No text was found.	Poor / None

Lay Health Workers

Table C.1.3 - Criterion Table for Sloan, N.L.; Community-Based Kangaroo Mother Care to Prevent Neonatal and Infant Mortality: A Randomized, Controlled Cluster Trial, Pediatrics 2008; 121; e1047[6]

Criterion	Example of text related to this criterion	
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	The training and intervention delivery processes that were used in the pilot study were only partially transferred to the full trial, with unplanned substitution of experienced trainers with individuals who were not trainers, less frequent contact between community workers and mothers in the last month of pregnancy, and unplanned emphasis on CKMC for LBW infants. The nested qualitative study found that >35% of CKMC women were erroneously taught that STS was to be provided to LBW or preterm infants rather than to all infants, and only 30% were correctly taught to hold all infants STS. Less than 40% of CKMC mothers were taught to provide CKMC to infants who were ill, and >25% were erroneously taught to breastfeed on schedule (not on demand). Women had numerous views about the number of hours and days they should provide STS, indicating that they received variable and frequently incorrect messages from the community workers and supervisors. *(Indicates implementation was not as successful as planned)*	Good

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	The Centers for Disease Control and Prevention (CDC) and the Pan American Health Organization (PAHO) have developed an inexpensive, rapidly implementable alternative for water quality improvement. This intervention consists of three elements: (1) point-of-use treatment of contaminated source water with disinfectant produced locally using appropriate technology; (2) safe storage of treated water; (3) community education. *(Indicates the intervention was developed externally)*	Fair

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	No text was found.	Poor None
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	No text was found.	Poor / None

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	No text was found.	Poor / None

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	Only mentions “community health workers” and mentions that they were already promoting chlorination in both study communities before the study itself.	Fair
b. Characteristics of the recipients	The population is described in detail on page 85 of the article along characteristics such as baseline water and hygiene practices. It covers 127 households with 791 persons, 50% of whom have <6 years of schooling and mean per capita income is $230. I don’t quote the entire thing since it’s a few paragraphs long.	Good
c. The setting	The setting and the characteristics are very similar if not identical in this article. The setting is two periurban communities of Montero, Bolivia, a city located in the subtropical eastern lowlands. Shallow uncovered household wells were the primary drinking water source for 87% of the households included.	Good
d. The mode of delivery (such as face-to-face)	The intervention is delivered face-to-face.“In July 1994, we interviewed the person responsible for handling water in the household, usually the female head of household or oldest daughter, about family socioeconomic and demographic characteristics, hygienic habits and water handling practices….”	Good

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
e. The intensity of the intervention/program (such as the contact time with participants)	This is not expressly mentioned and is hard to track from a lengthy description of the overall study design including differences in intensity across intervention and control households. On page 84 the article states (after discussing a baseline survey in July 1994, a baseline water test done on all households in August 1994, and mention of a public lottery to assign households to treatment and control groups), "From 22 to 25 August 1994, community health volunteers distributed one container of disinfectant and two special vessels to each intervention household and explained how to treat and store water with these products. Once a week, community health volunteers distributed containers with freshly prepared disinfectant to each intervention household, removed old containers, and used the labels on the special vessels to reinforce messages about proper use of the disinfectant and vessels and remind participants of different applications for treated water. Six visits at monthly intervals were made to all participating households from September 1994 to February 1995 to survey water-handling practices and to test stored and source water quality as described above. From 1 October 1994 to 28 February 1995, a specially-trained health worker made weekly visits to all households to obtain information about all household cases of diarrhoea, defined as >= 3 loose or watery stools in 24 h, with onset in the preceding 7 days."	Good
f. The duration (such as the number of sessions and their spacing interval over a given period)	This might be answered by e. above, but again not much detail on precisely how long and frequent the visits were, and by whom.	Fair
g. Adherence or fidelity to delivery protocols	No mention of this by community health volunteers tasked with implementing the intervention.	None

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	"This intervention is a promising way of providing microbiologically safe water in developing countries. While supplying piped, treated water to all households remains elusive for many communities, this point-of- use disinfection and safe water storage intervention can be rapidly disseminated, is inexpensive, simple to use, and adaptable to a variety of conditions. A similar water vessel can be manufactured in Bolivia at a cost of under US $4·00 each. The disinfectant can be produced in any community by inexpensive, solar- powered electrolysis of a salt water solution for as little as $0·05 per family per year. An earlier cost-effectiveness study estimated that this intervention would have no net cost to society if it decreased diarrhoea incidence by 20% or more. Ultimately, the utility of this intervention will be determined by its acceptability and sustainability in diverse populations. Social marketing will be an important component of efforts to enhance the intervention's acceptability and to ensure its sustainability through commercialization. An attractive aspect of this intervention is that it yields a product, the disinfectant, which can be marketed beyond the community as an alternative to boiling, which is expensive and time-consuming, and to solar disinfection, which is time-consuming and does not prevent recontamination….. Start-up costs for the production of the special vessels, disinfectant, and promotional materials, and for the establishment of distribution networks, will be substantial. The prospect of local management of the project and either full or partial cost recovery enhances the potential for success. Further ` real life ' investigations of this and similar strategies in other communities and at a larger implementation scale will define better the potential of this promising new strategy for waterborne disease prevention."	Fair

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	The only description of population needs was that diarrheal diseases, frequently transmitted by faecally-contaminated water, continue to be a major source of morbidity and mortality in developing countries. It later mentions that much of the local water is contaminated in the study area.	Fair
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	No text found.	Poor / None

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	There is great detail on a biologic/scientific level in terms of the chemical make-up of the chlorine solution or the style of buckets distributed, but there is only passing reference to “messages” given to households to reinforce the importance of water treatment without further detail on these informational or educational components.	Poor

Household Water Chlorination

Table C.2.1 - Criterion Table for Quick et al.; Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. 1999. Epidemiol. Infect. 122:83-90.[7]

Criterion	Example of text related to this criterion	Rating
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	No relevant text found, but need for further implementation in other settings to determine potential success of the intervention is mentioned.	Poor / None

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	The field work was done by Health Oriented Preventive Education (HOPE), a local non-governmental organisation that operates health clinics and undertakes community-based health and development initiatives in the area. *(Indicates the intervention was developed externally)*	Fair

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	No text was found.	Poor / None
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	No text was found.	Poor / None

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #4 Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	No text was found.	Poo/none

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	There is mention of the field workers: "Field workers, recruited from the study or nearby communities, were extensively trained in interviewing techniques, in data recording, in general approaches to community motivation and in specific techniques for promoting hand washing and drinking water treatment. The same field workers promoted regular use of the interventions and collected outcome data during their household visits."	Good
b. Characteristics of the recipients	Table 1 has summary statistics on the study sample divided by intervention status. Characteristics include average household size, number of children less than 5 and less than 2 years old, number of rooms in house, literacy of the mother, occupation of the father, and average expenditures on water per week, etc.	Good
c. The setting	The setting is described as "This study was conducted in adjoining multi-ethnic squatter settlements in central Karachi – Bhittaiabad, Bilal Colony, Mujahid Colony, Manzoor Colony and Zia Colony." Later it mentions "Field workers identified communities that typically received at least one hour of running water twice weekly, and had not received soap or water treatment in a previous study with HOPE.	Good
d. The mode of delivery (such as face-to-face)	The intervention is delivered face-to-face. "Field workers arranged neighbourhood meetings. They used slide shows, videotapes and pamphlets to illustrate health problems resulting from hand and water contamnation and to provide specific instructions on how to use the intervention assigned to that neighbourhood. Field workers, who spoke the first language of the study subjects, visited each participating household at least twice weekly. They encouraged use of the interventions, answered questions, and provided families with the consumable supplies necessary for ongoing use of the intervention".	Good

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
e. The intensity of the intervention/program (such as the contact time with participants)	This is not expressly mentioned in terms of length of visits, but on page 481 it mentions that visits to control households were generally shorter since no health education or encouragement for behavior change was provided. It writes, "Field workers visited participating households at least weekly, for 37 weeks (April 2003-December 2003), and asked the mother or other caregiver if the children had diarrhea (three or more loose stools within 24 h) in the preceding week, and, if so, for how many days. The mother was also asked about her own symptoms of diarrhea. Typically, field workers visited each household twice during the week to ensure that episodes of diarrhoea from both early and late in the week were recalled."	Good
f. The duration (such as the number of sessions and their spacing interval over a given period)	Again, this might be answered by e. above.	Good
g. Adherence or fidelity to delivery protocols	Delivery protocols in terms of whether the correct intervention was delivered including proper health messages is not addressed, but seemingly such issues would be discovered during checks to confirm diarrheal outcome measures: "Supervisors revisited 40% of homes each week and reviewed the history of diarrhoea among family members. The history recorded by the supervisor was compared to the history recorded by the field worker, and if there was a discrepancy, the fieldworker and supervisor revisited the house to clarify the difference.".	Fair
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	Costs are not calculated nor considered in detail. However, again this appears to be an earlier stage efficacy trial that argues for further study and such considerations to be taken into account.	Poor / None

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	The only description of population needs was that diarrheal diseases are a leading cause of childhood death in squatter settlements such as the ones they study.	Poor
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	No text found.	Poor / None

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	There is great detail on a biologic/scientific level in terms of the chemical products and soap distributed, but there is only passing reference to "messages" given to households to reinforce the importance of water treatment without further detail on these informational or educational components.	Poor

Household Water Chlorination

Table C.2.2 - Criterion Table for Luby et al.; Combining drinking water treatment and hand washing for diarrhea preventions, a cluster randomized controlled trial. 2006. Tropical Medicine and International Health, 2(4):479-489.[8]

Criterion	Example of text related to this criterion	Rating
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	Authors expressly write that outside of an RCT that included free supplies and twice weekly visits adherence to water treatment and hand washing would likely be much less. Say the next step is to implement them at larger scale and evaluate their practicality, uptake and effectiveness.	Poor

Household Water Chlorination

Table C.2.3 - Criterion Table for Crump et al.; Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhea in areas with turbid source water in rural western Kenya: cluster randomized controlled trial. 2005. BMJ, doi:10.1136/bmj.38512.618681.EO.[9]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	In response to the limitations, a new flocculant-disinfectant technology, that incorporates techniques used for municipal water purification, has been developed for treating water in the home. We hypothesised that this treatment could be useful in areas with turbid source water as the improvement in water clarity would encourage use. The lower chlorine demand associated with reduced turbidity and the removal of some chlorine-resistant organisms may provide a health advantage over sodium hypochlorite. We conducted a 20 week study to evaluate the efficacy of the flocculant-disinfectant in preventing diarrhoea in rural western Kenya, an area where source waters were both heavily faecally contaminated and highly turbid (100-1000 nephelometric turbidity units (NTU)).0 *(Basically, the intervention was developed externally.)*	Good
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully	No text is found.	Fair

Household Water Chlorination

Table C.2.3 - Criterion Table for Crump et al.; Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhea in areas with turbid source water in rural western Kenya: cluster randomized controlled trial. 2005. BMJ, doi:10.1136/bmj.38512.618681.EO.[9]

Criterion	Example of text related to this criterion	Rating
implemented in other countries."		
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	No text was found.	Poor / None
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	No text was found.	Poor/none
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:** a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	No description of the field workers.	Poor/none

Household Water Chlorination

Table C.2.3 - Criterion Table for Crump et al.; Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhea in areas with turbid source water in rural western Kenya: cluster randomized controlled trial. 2005. BMJ, doi:10.1136/bmj.38512.618681.EO.[9]

Criterion	Example of text related to this criterion	Rating
b. Characteristics of the recipients	The article cites three other articles for a more complete description of the population. Also, Table 1 has summary statistics on the study sample divided by intervention status. Characteristics include average household size, literacy status of household head, water source type (pond or river, etc.) and average baseline water quality measurements.	Good
c. The setting	The setting is described as "The study was conducted in 49 villages near Lake Victoria in Siaya and Bondo Districts, western Kenya. The demographic characteristics of people living in the area have been described elsewhere.An established clinic based surveillance system monitors the aetiology of diarrhoea among the population. Infant mortality is about 130 per 1000 inhabitants. Surface water used for drinking is typically obtained from ponds, rivers, and springs; it is regularly contaminated with both human and animal faeces. Water is typically carried in 20 l plastic drums and is stored in wide mouthed clay vessels holding.	Good
d. The mode of delivery (such as face-to-face)	Field workers visiting participating compounds weekly and used a standardized questionnaire to record the presence or absence of diarrhea and any deaths during the seven days since the last visit for each person.	Good
e. The intensity of the intervention/program (such as the contact time with participants)	The precise length of each visit is not made clear.	Poor/none
f. The duration (such as the number of sessions and their spacing interval over a given period)	There were different kinds of visits – weekly visits to ask about diarrhea, two longer surveys at weeks 5 and 15 to ask about attitudes towards the intervention, a baseline survey and unannounced visits every four weeks to collect water samples.	Good
g. Adherence or fidelity to delivery protocols	On behalf of the participants themselves, page 2: "**Compliance with intervention** Participants given flocculant-disinfectant retained the empty sachets after use. Each week field workers collected and counted empty sachets and replaced them. For participants using sodium hypochlorite, field workers collected and replaced bottles as needed each week. At the end of the study, partially used bottles were collected and weighed to determine the total use of sodium hypochlorite." *Nothing about adherence by field workers*.	Fair

Household Water Chlorination

Table C.2.3 - Criterion Table for Crump et al.; Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhea in areas with turbid source water in rural western Kenya: cluster randomized controlled trial. 2005. BMJ, doi:10.1136/bmj.38512.618681.EO.[9]

Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	Costs are not calculated nor considered in detail. Once again this appears to be an earlier stage efficacy trial that argues for further study.	Poor / None
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	Page 4, start of discussion section: In this setting where diarrhoea is a leading cause of childhood death and drinking water is highly turbid and contaminated with faeces, we found that children < 2 years from family compounds that treated their drinking water with flocculant-disinfectant had significantly less diarrhoea than compounds that used standard practices (control).	Poor / None
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned	No text found.	Poor / None

Household Water Chlorination

Table C.2.3 - Criterion Table for Crump et al.; Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhea in areas with turbid source water in rural western Kenya: cluster randomized controlled trial. 2005. BMJ, doi:10.1136/bmj.38512.618681.EO.[9]

in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."		
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	There is great detail on a biologic/scientific level in terms of the chemical products distributed, but not really any description if any information/education components were included.	Poor / None
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:**	No text was found.	Poor / None

Household Water Chlorination

Table C.2.3 - Criterion Table for Crump et al.; Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhea in areas with turbid source water in rural western Kenya: cluster randomized controlled trial. 2005. BMJ, doi:10.1136/bmj.38512.618681.EO.[9]

Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."		

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	The study was coordinated by the World Health Organization (WHO) with Principal Investigators at 5 universities in Africa. "The Kesho Bora study was a multicentre study conducted at five sites in three sub-Saharan African countries. Its multidisciplinary design required a large partnership of African research teams in the study sites as well as international research expertise in HIV and infectious diseases, obstetrics, paediatrics, nutrition, clinical trials implementation and analysis (Fig. 2). Several sponsors supported the project financially and technically. Such a large network necessitated strong coordination by the WHO Department of Reproductive Health and Research."	Good

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	"The Kesho Bora study was conceived before the recent rapid expansion of antiretroviral treatment (ART) programmes when antenatal care services often were unable to identify women requiring ART and even less able to provide access to ART. ART was known to decrease the risk of mother to child transmission of HIV."	Fair
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	"..the main goal of the Kesho Bora study was to optimize the use of ARVs during the antepartum, intrapartum, and postpartum periods for prevention of MTCT and for preserving maternal health." "Because MTCT risk and risk of maternal AIDS or death are strongly associated with maternal immunologic status, different ARV regimens were prescribed based on the mother's status (Table 1). "(All regimens in the Table include references).	Fair

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	"coordinated by the WHO Department of Reproductive Health and Research"	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	Characteristics of those delivering the intervention/program (such as a nurse or lay health worker) "Clinicians" "During the ARV initiation visit, study clinicians reviewed with each participant the drugs she would be receiving, their dosage, expected side effects and the optimal time of day for taking the drugs.	Poor / None Good
b. Characteristics of the recipients	Characteristics of the recipients "Participant eligibility criteria 1. Infected with HIV-1 2. Pregnant, with gestational age 20–32 weeks, with the exception of women with medically documented HIV stage 4 or CD4+ cell count b 200 cells/mm^3 who could be screened from gestational age 16 weeks. 3. Ability and willingness to give informed consent for screening (interview, physical examination, venipuncture for blood specimens, and estimation of gestational age) and home visits. 4. No evidence of clinically significant conditions (obstetric, cardiac, respiratory [including active tuberculosis], hepatic, gastrointestinal, endocrine, renal, haematologic, psychiatric, neurologic, or allergic) requiring care which may interfere with the study interventions. 5. Never enrolled in an HIV-vaccine trial. 6. No previous enrolment in the Kesho Bora study (for women who became pregnant again in the course of the study). 7. Not currently taking any ARV medications. 8. Capacity and willingness to participate in all follow-up visits, all clinical examinations and agreement for venipuncture for them and their babies.	

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
	9. Residing and planning to continue to reside in the study site catchment area until two years after delivery. 10. Willingness to receive and no contraindication to receive ARVs, i.e.: a. Severe anaemia (haemoglobin b 7 g/dl), b. Severe neutropenia (neutrophil count b750×10^6 cells/l) c. Blood alanine amino transferase N 5 times upper limit of normal (ULN) d. Amylase N 2 times ULN e. Blood creatinine N 3 times ULN f. Known allergy to one of the study ARVs or to benzodiazepines; g. Treatment with anticoagulants, benzodiazepines, rifampicin, magnesium sulphate, corticosteroids for more than 7 days at the time of planned enrolment."	
c. The setting	The setting "Site selection criteria 1. Ongoing MTCT prevention interventions (HIV counseling and testing during pregnancy, short-course ARV prophylaxis, and counseling regarding infant feeding options) through which HIV-1-infected, pregnant women could be recruited into the study; 2. Ability to enroll at least 120 eligible women per year (10 per month); 3. Capacity or potential capacity to follow enrolled women and their children adequately for 18–24 months after delivery with minimal loss of follow-up (no greater than 10% per year); 4. Significant proportion (≥ 50%) of HIV-1-infected mothers choosing to breastfeed their infants despite counseling on infant feeding options and the availability of free or low-cost infant formula; and 5. Services for long-term HIV care, including CD4+ cell count monitoring and ART when needed (either at study initiation or with reasonable expectation that access will be available within two years after study initiation)."	Good
d. The mode of delivery (such as face-to-face)	The mode of delivery (such as face-to-face) "face-to-face" "Follow-up visits included counseling, interviews and physical examinations."	Good

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
e. The intensity of the intervention/program (such as the contact time with participants)	The intensity of the intervention/program (such as the contact time with participants) "Participants had scheduled study visits weekly until eight weeks after delivery, monthly until 12 months after delivery and every three months thereafter."	Fair
f. The duration (such as the number of sessions and their spacing interval over a given period)	The duration (such as the number of sessions and their spacing interval over a given period) "Participants had scheduled study visits weekly until eight weeks after delivery, monthly until 12 months after delivery and every three months thereafter."	Fair
g. Adherence or fidelity to delivery protocols	Adherence or fidelity to delivery protocols "Women in Part IB received a short-course ARV prophylaxis regimen as per WHO recommendations, which consisted of 300 mg AZT taken by the mother twice daily starting from 34 to 36 weeks of pregnancy until the onset of labour, plus one 600 mg dose of AZT and one 200 mg dose of NVP at the onset of labour." "From 2007 following a change in the WHO guidelines, ARV prophylaxis was started in both arms in Part II from 28 weeks of pregnancy, the new recommended time for starting the short-course regimen because of greater effectiveness than the previously recommended start at 34–36 weeks. In addition AZT 300 mg with 3TC 150 mg twice daily for one week postpartum was added to reduce the risk of selection for NVP resistance in the mother. All protocol versions and amendments are summarized in Table 3." "Infant feeding counseling based on UNICEF/WHO training courses on Breastfeeding Management"	Fair

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	Not reported.	Poor / None
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	"At the time of study initiation, ART programmes were only beginning to be implemented. Study sites were purposely chosen in areas where programmes to increase long-term access to ART were already established, or in development, to ensure long-term access to HIV disease monitoring and treatment for participants after study completion."	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	"In Mombasa, for example, implementation within the public provincial hospital required a large and multidisciplinary team of dedicated research staff and part-time government employed health care providers. This partnership resulted in important exchange in resources, but also in logistic challenges, particularly due to high turn-over of government staff. The paperwork involved in clinical trials (approximately 60 different CRFs) was very time-consuming and resulted in large logistic challenges for clinicians and the data management team." "High mobility of many participants over the duration of the study caused difficulties in subject tracing, follow-up and retention in this rural-urban setting. Poverty of the rural participants presented an additional challenge in providing, for example, nutritional counseling to ensure adequate child growth and maternal nutrition." "The study enrolled participants at a lower rate than expected due to several factors. First, there were financial constraints. Funding was initially secured for only three sites (Bobo Dioulasso, Mombasa and Nairobi). It was decided to launch the project while trying to secure funds for two additional sites, with a backup plan to extend duration of recruitment in the original three sites if necessary. Funds for two planned sites in Rwanda and Tanzania were never identified, but funding for two South African sites was secured almost two years after enrolment of the first participant. Other factors negatively affecting recruitment included a delay in initiating the RCT by more than six months following the FDA advisory regarding NVP in women with CD4+ cell counts N250 cells/mm^3, and the lower than estimated prevalence of HIV-1-infected pregnant women in Bobo Dioulasso."	Good

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	"Children's HIV-1 infection status was assessed using a quantitative HIV-1 RNA real-time PCR assay (Generic HIV-1 Charge Virale, Biocentric, Bandol, France) in all sites except Nairobi where a qualitative HIV-1 DNA PCR assay (Amplicor HIV-1 DNA v1.5 assay, Roche) was initially used and infection status of all children considered positive subsequently confirmed using the quantitative real-time PCR assay." "All infants received a single dose of NVP (0.6 ml oral suspension, approximately 2 mg/kg body weight) within 72 h of birth. From 2007, one week of AZT (4 mg/kg twice daily) was added to reduce the risk of selection for NVP resistance in infected infants (Table 3)."	Fair

Preventing Mother-to-Child Transmission

Table C.3.1. - Criterion Table for Kesho Bora study group, 2010. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery, and breastfeeding for prevention of mother to children transmission of HIV-1: The Kesho Bora Multicentre Collaborate Study rational, design, and implementation challenges.[10]

Criterion	Example of text related to this criterion	Rating
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	"In each site, study implementation was monitored by dedicated quality assurance staff as well as by the WHO Kesho Bora Site Coordinator and an independent external Good Clinical Practice (GCP) monitor."	Fair

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	The only text found is: “The study is a retrospective analysis of mother and infant data from the hospital records of the perinatal and HIV PMTCT service of the Municipal Hospital Divina Providencia, a general population hospital situated in the urban area of Luanda, Angola.”	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	"Antiretroviral therapy is effective in reducing rates of mother-to child transmission of HIV to low levels in resource-limited contexts (REFs)"	Fair
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	"Women accessing the HIV PMTCT and perinatal care service at the Luanda Divina Providencia hospital are managed according to standardized procedures."	Poor / None
Criterion #4	Not described.	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Outer Setting: External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.		

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

<table>
<tr><th>Criterion</th><th>Example of text related to this criterion</th><th>Rating</th></tr>
<tr><td>Criterion #5
Intervention Characteristics:
Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4]
The detailed description should include:</td><td></td><td></td></tr>
<tr><td>a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)</td><td>Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)
“Clinicians” - presumably MDs + nurses at the hospital</td><td>Poor / None</td></tr>
<tr><td>b. Characteristics of the recipients</td><td>Characteristics of the recipients
<table>
<tr><td></td><td>N</td><td>Mean</td></tr>
<tr><td>Age (years)</td><td>104</td><td>29.2</td></tr>
<tr><td>Body mass index (kg/m^2)</td><td>102</td><td>23.1</td></tr>
<tr><td>Number of previous pregnancies</td><td>104</td><td>2.9</td></tr>
<tr><td>Number of previous live births</td><td>104</td><td>2.5</td></tr>
</table></td><td>Fair</td></tr>
<tr><td>c. The setting</td><td>The setting
“The study is a retrospective analysis of mother and infant data from the hospital records of the perinatal and HIV PMTCT service of the Municipal Hospital Divina Providencia, a general population hospital situated in the urban area of Luanda, Angola.”</td><td>Fair</td></tr>
<tr><td>d. The mode of delivery (such as face-to-face)</td><td>The mode of delivery (such as face-to-face)
Face-to-face</td><td>Fair</td></tr>
<tr><td>e. The intensity of the intervention/program (such as the contact time with participants)</td><td>The intensity of the intervention/program (such as the contact time with participants)
“At least one prenatal or perinatal service access”</td><td>Poor / None</td></tr>
<tr><td>f. The duration (such as the number of sessions and their spacing interval over a given period)</td><td>The duration (such as the number of sessions and their spacing interval over a given period)
“At least one prenatal or perinatal service access”</td><td>Poor / None</td></tr>
<tr><td>g. Adherence or fidelity to delivery protocols</td><td>Adherence or fidelity to delivery protocols
No information on adherence or fidelity. “Women accessing the HIV PMTCT and perinatal care service at the Luanda Divina Providencia hospital are managed according to standardized procedures.”</td><td>Poor / None</td></tr>
</table>

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	Not reported.	Poor / None
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	Not reported.	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	Not reported.	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	Not described.	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.2 - Criterion Table for Lussiana, 2012. Effectiveness of a Prevention of Mother-to-Child HIV Transmission Programme in an Urban Hospital in Angola.[11]

Criterion	Example of text related to this criterion	Rating
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	Retrospective analysis of patient records for medication adherence and follow-up attendance.	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #1 **Intervention Characteristics:** Intervention/Program source (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** Is the intervention/program externally or internally developed? An intervention/program may be internally developed as a good idea, a solution to a problem, or other grass roots effort, or may be developed by an external entity (such as a foundation or a NGO). Interventions or programs that arise internally from the populations who will be impacted are sometimes more sustainable than externally developed programs dependent on external funding. The perceived legitimacy of the source may also influence implementation.	"Baylor College of Medicine Children's Foundation Malawi, in collaboration with the Malawi Ministry of Health (MOH), initiated a pilot community-based intervention" "Prior to the programme intervention, consultative meetings were conducted with community leaders. CHWs conducted daily education sessions in the health centres and held ongoing sensitization meetings in the community. The main focus of education was promoting the utilization of PMTCT, EID and paediatric HIV treatment services."	Good

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #2 **Intervention Characteristics:** A description of why the intervention was hypothesized to have an impact on the outcome, according to theory. (From CReDECI, Mohler 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** The theoretical basis of the intervention should be clearly stated. This includes the theory on which the intervention is founded as well as, if available, empirical evidence from studies in different settings or countries. For example, "The implementation was based on Rogers' Diffusion of Innovation theory, which posits 5 factors of innovation that influence a decision to adopt or reject an innovation: relative advantage, compatibility, complexity or simplicity, trialability, observability. A similar intervention, also based on Rogers' Diffusion of Innovation theory, was successfully implemented in other countries."	"The intervention was designed to create a new paradigm in PMTCT service delivery and end the compartmentalization of services into distinct PMTCT, EID and paediatric HIV subunits. Tingathe CHWs ensured longitudinal care throughout the full PMTCT cascade, starting with diagnosis of the mother at antenatal care (ANC) and ending with final diagnosis and treatment of the infant."	Fair
Criterion #3 **Intervention Characteristics:** Rationale for the aim/essential functions of the intervention/program's components, including the evidence whether the components are appropriate for achieving this goal. This differs from the need to articulate the theory behind the intervention in that the theory posits the general principles (such as Rogers Diffusion of Innovation) while this item is about specific components of the intervention and the effects of the component on specific targets. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4]	"Programs in high burden countries, including Malawi, often subdivide aspects of this cascade into separate PMTCT (vertical transmission), antiretroviral therapy (ART), early infant diagnosis (EID) and paediatric HIV programs, frequently with different providers and service locations for each component. Resulting poor utilization of available services, lack of coordination between providers and high rates of loss to follow-up have led to persistent high infection rates in exposed children. It has been shown that, even with highly efficacious combination antiretroviral interventions, only marginal reductions in childhood HIV infections can be achieved without improved retention of pregnant mothers and infants within the PMTCT cascade of services. Task shifting with the use of community health workers (CHWs) has been suggested as one strategy to address these challenges within resource-limited settings. (REFs)"	Fair

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #4 **Outer Setting:** External policies and incentives (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** How does the health service, intervention, or program relate to country and global health goals? Is the program part of a larger strategy? If so how is it strategically aligned? A country's health policies may influence the implementation of a particular intervention or program.	"In 2011, UNAIDS announced a call to eliminate new paediatric HIV infections among children by 2015."	Fair

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #5 **Intervention Characteristics:** Detailed description of the intervention/program (From WIDER as described in Michie, 2009)[4] **The detailed description should include:**		
a. Characteristics of those delivering the intervention/program (such as a nurse or lay health worker)	"Criterion for CHW selection included living within the community, completion of primary schooling and ability to read and write in English and Chichewa, ability to ride a bicycle and HIV-infected or affected. Both men and women were recruited."	Good
b. Characteristics of the recipients	"Pregnant women living with HIV and their exposed infants"	Fair
c. The setting	"The Tingathe-PMTCT pilot programme took place in Area 25 and Kawale, two large peri-urban communities in Lilongwe. The estimated population is 310,000 people, with 15,000 deliveries/year, 2000 HIV-exposed infants delivered/year and 12% adult HIV prevalence."	Good
d. The mode of delivery (such as face-to-face)	Face to face. "CHWs ensured that mother-infant pairs received all necessary PMTCT services. They followed their clients at their homes and at health centres, from initial diagnosis up until confirmation of definitive HIV-uninfected status after cessation of breastfeeding or successful ART initiation for HIV-infected infants."	Good
e. The intensity of the intervention/program (such as the contact time with participants)	Varies	Fair
f. The duration (such as the number of sessions and their spacing interval over a given period)	"From initial diagnosis up until confirmation of definitive HIV-uninfected status after cessation of breastfeeding or successful ART initiation for HIV-infected infants."	Good
g. Adherence or fidelity to delivery protocols	"All PMTCT clinical care was provided in accordance with MOH and WHO guidelines."	Poor / None

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #6 **Intervention Characteristics:** Costs of the intervention and costs associated with implementing the intervention (From CFIR, Damschroder, 2009; CReDECI, Mohler, 2012)[2,3] **Explanation/Example:** The cost of the intervention and implementation can influence the adoption and sustainability; interventions maybe more difficult to sustain if they were supported as part of a research study.	Not reported.	Poor / None
Criterion #7 **Population needs** (From CFIR, Damschroder, 2009)[2] **Explanation/Example:** The extent to which population needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritized. This could include population-based data on causes of morbidity and mortality, political or cultural barriers or facilitators, and/or more locally focused data about local needs, barriers or facilitators.	"We used three sources for preintervention data. The first was a published report of maternal and infant utilization of PMTCT, EID and paediatric HIV services at five sites (including our two intervention sites) within Lilongwe between 2004 and 2008 (REF). This source contained preintervention comparison data for PMTCT prophylaxis, infant PCRs and ART initiation for HIV-infected infants. For information not included in this report, we used the 2004 Malawi Demographic and Health Survey, which provided national statistics for numbers of women accessing ANC, location of delivery and infant feeding choice after birth."	Good

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #8 **Process of implementation:** Description of facilitators or barriers which have influenced the intervention or program's implementation (see #10) revealed by a process assessment. In contrast to the criterion #7 above which assesses barriers and facilitators as inputs to developing the intervention strategy, this criterion assesses the actual barriers and facilitators identified during and after the implementation. (From CReDECI, Mohler, 2012; also mentioned in Michie, 2009)[3,4] **Explanation/Example:** "The attitudes of the nursing home managers turned out to be an important factor supporting or impeding the success of the intervention's implementation. The more the managers agreed with the interventions' aim, the better the nursing staff felt supported."	"The strongest predictors of successful completion of the PMTCT cascade were enrolment in the third trimester (OR, 0.37; 95% CI, 0.24 to 0.58), having newly diagnosed HIV infection (OR, 0.50; 95% CI, 0.33 to 0.75) and having a partner who was not involved (OR, 0.43; 95% CI, 0.24 to 0.78)."	Good

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #9 **Description of materials:** Description of all materials or tools used for the implementation (From CReDECI, Mohler, 2012)[3] **Explanation/Example:** "The primary enablers of behaviour change were paid community-based health workers, who were recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. They received a combination of classroombased and apprentice ship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building. After training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 trainees, for an additional week before final selection was made."	Community Health Worker curriculum: 1. Basics of HIV/AIDS 2. PMTCT: what are the steps and how to promote utilization of services 3. Caring for the exposed infant: importance of early infant diagnosis and cotrimoxazole prophylaxis 4. Diagnosing HIV infection 5. Nutrition: exclusive breast feeding, complementary feeding, and malnutrition screening 6. Children with HIV: identification, care and treatment 7. Anti-retroviral therapy and adherence counseling 8. Reducing stigma and discrimination 9. Counseling and community mobilization and education skills 10. Conducting the patient home visit"	Good

Preventing Mother-to-Child Transmission

Table C.3.3 - Criterion Table for Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, Chiao EY, Schutze GE, Kline MW. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID 22789644.[12]

Criterion	Example of text related to this criterion	Rating
Criterion #10 **Process of Implementation:** Description of an assessment of the implementation process (From CReDECI, Mohler 2012)[3] **Explanation/Example:** Process assessment is a prerequisite for determining the success of the intervention's implementation and should be an integral part of an assessment of the intervention's effect. For example, "To gain insight into the dissemination and the delivery of the intervention and to draw conclusions about potential barriers and facilitators to implementing the intervention in other settings, data on the implementation process were collected alongside the randomized-controlled trial. Therefore, we assessed the quality of delivery of the interventional components (observed by members of the research team not involved in the delivery of the intervention) and the adherence to study protocol (number and type of deviations from the protocol, using a pilot-tested standardized form). We also analyzed barriers and facilitators for the delivery of intervention's components (focus group interviews with intervention participants)."	"An individual patient mastercard was used to facilitate patient case management, and a patient register was used to monitor CHW activities. The mother-infant mastercard was opened on programme entry, updated after every visit and key data entered into registers weekly. Information from registers was entered into a Microsoft Access database bimonthly."	Fair

References

1. Kumar V, Mohanty S, Kumar A, et al. Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. Lancet. 2008;372(9644):1151-62. PMID: 18926277.
2. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:50. PMID: 19664226.
3. Mohler R, Bartoszek G, Kopke S, et al. Proposed criteria for reporting the development and evaluation of complex interventions in healthcare (CReDECI): guideline development. Int J Nurs Stud. 2012;49(1):40-6. PMID: 21924424.
4. Michie S, Fixsen D, Grimshaw JM, et al. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci. 2009;4:40. PMID: 19607700.
5. Kouyate B, Some F, Jahn A, et al. Process and effects of a community intervention on malaria in rural Burkina Faso: randomized controlled trial. Malar J. 2008;7:50. PMID: 18364043.
6. Sloan NL, Ahmed S, Mitra SN, et al. Community-based kangaroo mother care to prevent neonatal and infant mortality: a randomized, controlled cluster trial. Pediatrics. 2008;121(5):e1047-59. PMID: 18450847.
7. Quick RE, Venczel LV, Mintz ED, et al. Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. Epidemiol Infect. 1999;122(1):83-90. PMID: 10098789.
8. Luby SP, Agboatwalla M, Painter J, et al. Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomised controlled trial. Trop Med Int Health. 2006;11(4):479-89. PMID: 16553931.
9. Crump JA, Otieno PO, Slutsker L, et al. Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhoea in areas with turbid source water in rural western Kenya: cluster randomised controlled trial. BMJ. 2005;331(7515):478. PMID: 16046440.
10. Safety and effectiveness of antiretroviral drugs during pregnancy, delivery and breastfeeding for prevention of mother-to-child transmission of HIV-1: the Kesho Bora Multicentre Collaborative Study rationale, design, and implementation challenges. Contemp Clin Trials. 2011;32(1):74-85. PMID: 20854932.
11. Lussiana C, Clemente SV, Ghelardi A, et al. Effectiveness of a prevention of mother-to-child HIV transmission programme in an urban hospital in Angola. PLoS One. 2012;7(4):e36381. PMID: 22558455.
12. Kim MH, Ahmed S, Buck WC, et al. The Tingathe programme: a pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. J Int AIDS Soc. 2012;15(4):1-11. PMID: 22789644.