

**Review Article**

# Making a Case for Measuring Implementation Strength of Programs

**Running Title:** Measuring Implementation Strength of Programs

Sushama K.Sreedhara <sup>1,2</sup>, Diwakar Mohan <sup>1</sup>, Melissa A.Marx <sup>1</sup>

<sup>1</sup>International Institute for Programs, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>2</sup>Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

**Corresponding Author:**

Dr Diwakar Mohan. Email: [awid79@gmail.com](mailto:awid79@gmail.com)

## 1. Introduction

The healthcare market is broadly categorized into a supply side and a demand side. The demand side includes individuals and households receiving care and has been studied quite extensively with respect to their responses to price, quality, type, waiting time, and distance from a healthcare provider. The supply side pertains to all the providers and sources who deliver and distribute health care. Health care providers are legal or administrative units who produce healthcare services, and they range from simple entities like individual practitioners, to complex groups like large hospitals. In low income countries, knowledge about this supply side market is usually inadequate.[1] From the implementers perspective, optimizing the supply aspects of healthcare operations efficiently would be more beneficial than the demand side. Programs can induce demand in several ways including geographic accessibility to health services, organizing demand

generation activities (for e.g. vaccination awareness drives), improving quality of care and providing affordable health services. Implementers control and plan for activities, proactive or reactionary, managing human resources to deliver healthcare, providing equipment and supplies (both medical, and non-medical), training health care workers, monitoring programs, etc. Considering that billions of dollars are spent on healthcare programs to strengthen the supply side, it is important that appropriate tools are used to monitor and evaluate these programs.

## 2. Logic Model for Program Evaluation

Program managers have been using logic model for more than 20 years, to describe the underlying framework of their programs. It helps in designing programs, sharing ideas, building teams, recognizing assumptions made, communicating progress, pinpointing

high impact projects, and so on. They are used to design evaluation plans, monitor and evaluate progress of the projects, collect relevant data, guide evaluators to assess the program more easily, measure the level of impact, and communicate results.[2,3] Logic model which is also known as “Chains of reasoning”, “Theory of action”, and “Logical Framework”, is a graphical diagram that describes the assumptions any program makes to lead to a well-defined outcome.[2] It includes stages like inputs, processes, outputs, short, intermediate and long-term outcomes, and impacts. It demonstrates sequences of cause-effect relations and is part of a systems approach to describe the path to an expected outcome. [2-4] It can, pictorially, demonstrate how a program feature (“process” of training healthcare workers) affect another (“output” like quality of care).[2-4] The Common Evaluation Framework is an example of the comprehensive approach to monitoring and evaluation of public health programs[5] (fig 1).

A key component of the logic model is the concept of Implementation Strength (IS). This concept paper discusses IS as a measure of structural quality and provides guidance for developing assessments of IS.

### 3. Implementation Strength

Implementation refers to “specified set of activities designed to put into practice an activity or program of known dimensions”.[6] There are several aspects of program implementation that have been studied. The degree to which some specified activities are implemented as planned is referred to as “integrity”, “fidelity” or “adherence”.[7-9] Implementation strength is intended to measure the amount of a program that is delivered (from the supply side), rather than how

much the program is received (utilization or coverage).[10]

Implementation strength can be used to guide efforts to enhance the effectiveness of programs and assess the intensity and extent of scale up, so that they reach the population intended. Implementation strength is used to assess the relationship between the program (“dose”) and expected outcomes, such as higher intervention coverage and lower under five mortalities (“response”).[10] Implementation strength also includes some aspects of what has traditionally treated as ‘structural’ or ‘readiness’ domain of quality of care.

One of the ways to understand the relationship between public health gains and programs, is to measure the intensity with which intervention packages are delivered. These measures help in attributing changes in a dependent variable (outcome) through manipulation of an independent variable (implementation).[11] There are several reasons to evaluate strategies for improving public health, including impact evaluations for attribution, course correction for making the program more effective, providing accountability to donor organizations, and assessing the internal and external validity of interventions.[9] Without measuring implementation data, accurate interpretations of outcomes/results obtained cannot be made, because it is hard to assess if failures were due to insufficient or incomplete delivery of services, or due to poor planning/ conceptualization of the program.

Designs like randomized controlled trials(RCT) and quasi-experimental studies which have control and study areas are not always feasible under

conditions where programs are being scaled up uniformly across all the regions of a country.[11] It is essential to recognize what interventions can be delivered at scale, how they can be tuned to work in new settings, and what the public health benefits of scaling up the efforts are.[11] A practical alternative to experimental designs would be to measure the strength or intensity with which packages of interventions are delivered upon roll out, so that the association between public health benefits and implementation can be better understood.[8]

Traditional program evaluation designs comparing intervention and comparison areas are less common, as programs are being rolled out in many regions simultaneously by one or more implementers.[11] RCTs are most suited to establish the efficacy of new behavioral or biological interventions.[11,12] Even though they are considered gold standard, they are hard to conduct on large scale basis in a real world setting.[12] Implementation of the same efficacious intervention tends to be less intense and more variable in real world conditions as compared to efficacy trials.

Since evaluators often cannot control when, where and how quickly the programs are scaled up by implementers, observational designs are necessary.[11,12] Hence, correct attribution of outcomes and impacts to program may be difficult. In such scenarios, knowing the implementation “dose”, outcome and impact changes at different administrative levels, can provide more conviction in the attribution pathway between program and impact. Measuring implementation strength, can help in course correction of programs if they aren't true to their

plans, understand why programs succeed or fail, anticipate outcomes of future programs, and help improve progress towards specific outcomes and intervention strategies.[13]

Implementation strength assessments can be performed using primary data collected explicitly for the purposes of the evaluation or secondary data collected for monitoring and documentation purposes. Assessments may be cross-sectional and performed at a single time point (snapshots) or may end up as longitudinal assessment, sharing near real-time feedback to researchers and staff at various stages of the implementation of a program. Having said that, methodological guidance is necessary to ensure that these assessments are developed correctly.

#### **4. Domains of Implementation Strength**

The various activities of program implementation can be broadly categorized into domains based on previous assessment tools like the Service Provision Assessments (SPA) and literature on process evaluation. As an example, the key domains emerging for Integrated Community Case Management programs (iCCM) for treatment of sick children, include training, supervision, service availability and deployment, provisioning of drugs and supplies, and routine monitoring. Programs may have different domains of focus and IS assessments need to be tailored to the nature of the program under evaluation.

The definition of domains helps with the identification of appropriate indicators to measure them. By listing indicators from the domains of IS, implementers will be able to explicitly

and exhaustively list the processes necessary for the successful implementation of any program. Additionally, since they have ready access to these indicators it enables them to focus on keeping track of the implementation status of their program, get real time feedback, and perform course correction, to be able to implement programs more effectively. There are several considerations for what constitutes a good indicator: (1) Indicators need to be SMART (Specific, Measurable, Attributable, Realistic and Time-Oriented). In essence, they will measure only the intended outcome and nothing else, they are clear, unambiguous and practical, they can help assign credit to projects for achieving the intended outcome, they must be reasonably aware of resources available to collect the data, and they must be time-sensitive.[15] (2) Applicability to a variety of contexts: Given the diverse nature of programs in terms of organizational sub-units, employment hierarchy and variability among

implemented program components, it is essential that these indicators be invariant to all the differences and be widely applicable.[10] (3) Validity, Reliability and Replicability: Indicators should be valid in the sense that they should accurately measure what they purport to measure. Upon repeated measurements, reliable indicators will return comparable results. The data returned by these indicators need also meet standards of replicability, in differing contexts. (4) Complementary relationship with existing reporting requirements: These indicators need to be aligned, but complementary to the reporting requirements of participating local and national governments or donors. However, like (2) above, local differences need to be adhered to, so that the indicators are valid.[10] Table 1 shows a sample of indicators for the evaluation of iCCM programs including details like the numerator and denominator, data collection frequency, and possible data source

<b>Table 1</b>					
<b>DOMAIN: SERVICE AVAILABILITY</b>					
<b>No</b>	<b>INDICATOR</b>	<b>NUMERATOR</b>	<b>DENOMINATOR</b>	<b>SOURCE OF DATA</b>	<b>FREQUENCY</b>
1	Percentage of communities with access to iCCM <sup>a</sup> services (Community or village or any other unit as defined by each country) [16, 18, 19]	Number of communities with active CHWs <sup>b</sup> in iCCM <sup>a</sup> services	Total number of target communities	Program Records	Yearly
<b>DOMAIN: PROVISIONS (drugs &amp; other supplies)</b>					
2	Percentage of CHWs <sup>b</sup> that had no stock out of iCCM <sup>a</sup> drugs/supplies <sup>c</sup> for any duration in the past 3 months (disaggregate by different drugs/supplies defined by country guideline) [20, 16, 21, 22, 23]	Number of CHWs <sup>b</sup> that had no stock out of iCCM <sup>a</sup> drugs/supplies <sup>c</sup> for any duration in the past 3 months	Total number of CHWs <sup>b</sup> who provide iCCM <sup>a</sup> services	Provider Assessment	Quarterly

3	Percentage of supply points <sup>d</sup> that had no stock out of iCCM <sup>a</sup> drugs/supplies <sup>c</sup> for any duration in the past 3 months (disaggregate by different supply points & list of drugs/supplies defined by country guideline)[20,16,21,22]	Number of supply points <sup>d</sup> that had no stock out of iCCM <sup>a</sup> drugs/supplies <sup>c</sup> for any duration in the past 3 months	Total number of supply points <sup>d</sup>	Supply point audit, Stock records	Quarterly
<b>DOMAIN: MONITORING</b>					
4	Percentage of health facilities implementing iCCM <sup>a</sup> that conducts quarterly meetings to go over data[24]	Number of health facilities implementing iCCM <sup>a</sup> that conducts quarterly meetings to go over data	Total number of health facilities implementing iCCM <sup>a</sup>	Program records, Provider interview	Quarterly
<b>DOMAIN: DEPLOYMENT</b>					
5	Ratio of CHWs <sup>b</sup> to supervisors for iCCM <sup>a</sup> at a target geographical area[20,21]	Number of CHWs <sup>b</sup> providing iCCM <sup>a</sup> services at a target geographical area	Number of supervisors for iCCM <sup>a</sup> services at a target geographical area	Program Records, Provider Assessment	Yearly
<b>DOMAIN: TRAINING</b>					
6	Percentage of supervisors ever trained in iCCM <sup>a</sup> [25]	Number of supervisors ever trained in iCCM <sup>a</sup>	Total number of supervisors	Program Records, Supervisor interviews	Yearly
<b>DOMAIN: SUPERVISION</b>					
7	Percentage of health facilities which received at least one supervision on iCCM <sup>a</sup> services in the last 6 months[17,26,27]	Number of health facilities which received at least one supervision on iCCM <sup>a</sup> services in the last 3 months;	Total number of health facilities providing iCCM <sup>a</sup> services	Facility Audit, Provider Assessment	Yearly
<p>a: Integrated Community Case Management  b: Community Health Workers  c: Example of Supplies (Forms: sick child, referral, counter-referral, case management charts/algorithm, patient register, danger sign charts, counseling card, supervisory checklist form, thermometer, timers, watches, spoon, cups, liter measures, access to drinking water, Mid Upper Arm Circumference (MUAC) tape, RDTs). Examples of drugs are Amoxicillin, ORS, Zinc, ACTs.  d: Supply point is a place where CHW can go to restock her supplies, it can be a health facility (which can be called as health post, health center) or pharmacy or other retail shops.</p>					

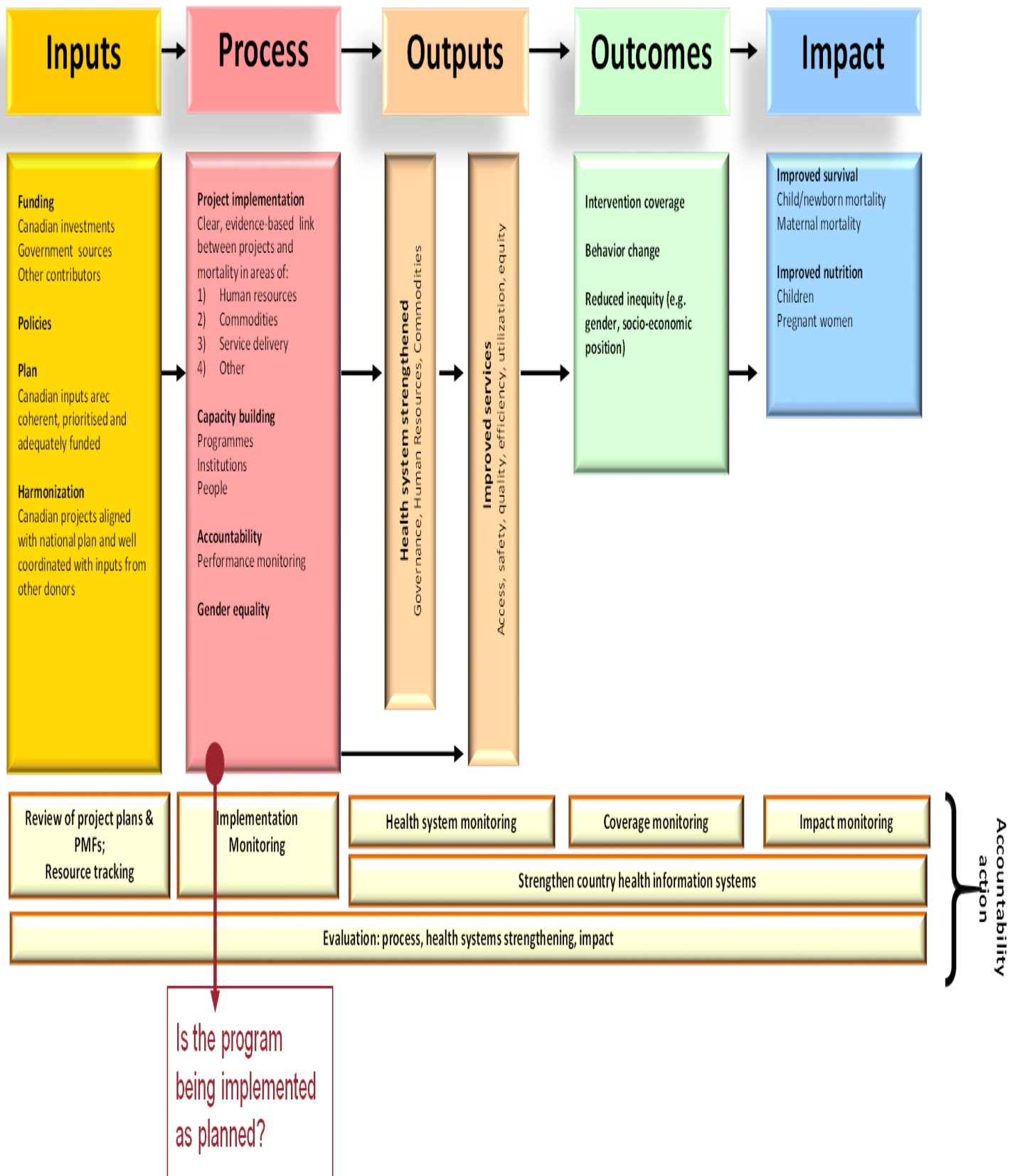


Figure 1 - Common Evaluation framework for programs

A stepwise approach to thinking about IS assessments is outlined below:

*(1) Logic/Impact model development:*

An appropriate logic/impact model of the intervention is needed as a precursor to the measurement of implementation strength. Any project can use an approximate separation of inputs and processes that are 'implementer-controlled', from the effects of these on target populations (as outputs, outcomes and impact).[8]

*(2) Understanding of the delivery platforms:* Identification of the program activities necessary to assess implementation.

*(3) Building consensus on appropriate indicators:* Literature reviews and discussions with experts in the field help to build consensus on appropriate indicators to measure implementation strength of the program planned to be evaluated.

*(4) Develop and design appropriate tools:* Appropriate survey tools need to be developed to assure that the right questions are asked, so that the indicators can be deduced correctly.

*(5) Implement data collection using appropriate sources:* An appropriate data collection methodology is needed to cover all the dimensions of implementations to be measured. Two considerations in the choice of data sources to note are the administrative level: household, district etc, and the time of measurement: multiple independent time points, baseline and end line only etc. Some examples of data sources include quality assurance data, monitoring and evaluation data, training reports, provider surveys, and program records.[8]

*(6) Analysis and presentation:* Implementation strength data can be analyzed in multiple ways. Categorical analysis creates distinct groups with varying levels of implementation and measures their differences to assess strength. A continuous measurement

could use percentages to assess levels of dosage or fidelity against a pre-determined threshold.[9] Sometimes it may be necessary to combine data from different domains of implementation strength into a single measure. This necessitates the consideration of weighting, either implicitly or explicitly. Different approaches for data reduction like principal component analysis, or review and consensus could be applied. Irrespective of how this is done, basic analysis like calibration and cross-validation would be necessary to validate the measures.[8] How these measures are used in analysis needs to be documented, ideally a priori, in a statistical analysis plan.[8]

*Real Accountability, Data Analysis for Results (RADAR)*

A comprehensive set of tools and aids are under development as part of the RADAR project implemented by the International Institute for Programs (IIP) at the Johns Hopkins Bloomberg School of Public Health with funding from the Government of Canada. One set of tools are those for measuring the implementation strength of a variety of programs. As part of the tool development process, the RADAR project plans to conduct implementation strength assessments in different Sub-Saharan African countries. Resources to aid in the impact evaluations including implementation strength assessments are available on the RADAR website (<https://www.jhsph.edu/research/centers-and-institutes/institute-for-international-programs/current-projects/RADAR/index.html>). Any queries regarding the use of resources may be addressed to the corresponding author.

## 6. Conclusion

Implementation strength assessments should be part of the toolkit for evaluation of programs in global health. From the perspective of donors and policy makers, assessing implementation strength would be useful to make decisions about their involvement in programs.

## 7. Funding

The authors are supported by funds from the RADAR project awarded to the Johns Hopkins Bloomberg School of Public Health from the Government of Canada. The authors declare no conflicts of interest.

## References

1. Berman, P. Understanding the supply side: a conceptual framework for describing and analyzing the provision of health care services with an application to Egypt. International Health Systems Group, Harvard School of Public Health. (1999)
2. McLaughlin, J. A., & Jordan, G. B. Logic models: a tool for telling your program's performance story. *Evaluation and program planning*, (1999); 22(1), 65-72.
3. Lawton, B., Brandon, P. R., Cicchinelli, L., & Kekahio, W. Logic Models: A Tool for Designing and Monitoring Program Evaluations. REL 2014-007. Regional Educational Laboratory Pacific.
4. Paul, F. M. The Logic Model for Program Planning and Evaluation (1997).
5. Bryce, J., Victora, C. G., Boerma, T., Peters, D. H., & Black, R. E. Evaluating the scale-up for maternal and child survival: a common framework. *International Health*, (2011);3(3), 139-146.
6. National Implementation Research Network (n.d.) Implementation defined. Retrieved 1 March 2017, from <http://nirn.fpg.unc.edu/learn-implementation/implementation-defined>
7. Dane, A.V. & Schnieder, B.H. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev*, (1998);18, 23-45.
8. Hargreaves, J. R., Goodman, C., Davey, C., Willey, B. A., Avan, B. I., & Armstrong, J. R. M. Measuring implementation strength: lessons from the evaluation of public health strategies in low- and middle-income settings,(2016);1-8. doi:10.1093/heapol/czw001
9. Durlak, J. A., & DuPre, E. P. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American journal of community psychology*, (2008); 41(3-4), 327.
10. Oliphant, N., Diaz, T., Daelmans, B., Psi, B., IIP-JHU, J. B., Public, L. D, LSHTM, J. S. Measuring the strength of implementation of community case management of childhood illness within the Catalytic Initiative to Save a Million



- Lives, (2011) ;(March), 2–3.
11. Victora, C. G., Black, R. E., Boerma, J. T., & Bryce, J. Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. *The Lancet*, (2011);377(9759), 85-95.
  12. Habicht, J. P., Victora, C. G., & Vaughan, J. P. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International journal of epidemiology*, (1999);28(1), 10-18.
  13. London School of Hygiene and Tropical Medicine and IDEAS (2012). Measuring Implementation Strength. Retrieved 20 January 2017, from [http://ideas.lshtm.ac.uk/sites/ideas.lshtm.ac.uk/files/Report\\_implementation\\_strength\\_Final\\_0.pdf](http://ideas.lshtm.ac.uk/sites/ideas.lshtm.ac.uk/files/Report_implementation_strength_Final_0.pdf)
  14. Wilson, S. J., Lipsey, M. W., & Derzon, J. H. The effects of school-based intervention programs on aggressive behavior: a meta-analysis (2003).
  15. Operational Policy and Quality Department: World Bank (2013). Results Framework and M&E Guidance Note. Retrieved 5 January 2017, from [http://siteresources.worldbank.org/PROJECTS/Resources/40940-1365611011935/Guidance\\_Note\\_Results\\_and\\_M&E.pdf](http://siteresources.worldbank.org/PROJECTS/Resources/40940-1365611011935/Guidance_Note_Results_and_M&E.pdf)
  16. CORE Group, Save the Children, BASICS and MCHIP (2012). Community Case Management Essentials: Treating Common Childhood Illnesses in the Community. A Guide for Program Managers. Retrieved 6 November 2016, from <http://www.coregroup.org/storage/documents/CCM/CCMEssentialsGuide/ccmbook2012-online.pdf>
  17. Miller, N. P., Amouzou, A., Bryce, J., Victora, C., Hazel, E., & Black, R. E. (2013). Assessment of iCCM implementation strength and quality of care in Oromia, Ethiopia. Baltimore, USA and Addis Ababa, Ethiopia: Institute for International Programs, Johns Hopkins Bloomberg School of Public Health.
  18. Kasungami, D., & Guenther, T. (2016). Overview of recommended indicators for routine monitoring of iCCM. Retrieved 28 January 2017, from <http://www.coregroup.org/storage/documents/iCCMPresentation.pdf>
  19. Heidkamp, R., Hazel, E., Nsona, H., Mleme, T., Jamali, A., & Bryce, J. Measuring implementation strength for integrated community case management in Malawi: results from a national cell phone census. *The American journal of tropical medicine and hygiene*, (2015);14-0797.
  20. USAID and MCHIP (2013). Indicator guide. Monitoring and evaluating integrated community case management, (Report). Retrieved 6 November 2016, from [http://pdf.usaid.gov/pdf\\_docs/PAA00JS1D.pdf](http://pdf.usaid.gov/pdf_docs/PAA00JS1D.pdf)
  21. Robertson, T., Kasungami, D., Guenther, T., & Hazel, E. (2016). Monitoring iCCM: a

- feasibility study of the indicator guide for monitoring and evaluating integrated community case management. Health policy and planning, czv129.
22. Diaz, T., Guenther, T., Oliphant, N. P., & Muñiz, M. A proposed model to conduct process and outcome evaluations and implementation research of child health programs in Africa using integrated community case management as an example. *Journal of global health*, (2014);4(2).
  23. Hazel, E., Amouzou, A., Park, L., Banda, B., Chimuna, T., Guenther, T., & Bryce, J. Real-time assessments of the strength of program implementation for community case management of childhood illness: validation of a mobile phone-based method in Malawi. *The American journal of tropical medicine and hygiene*, (2015);92(3), 660-665.
  24. UNICEF, World Health Organization and Ministry of Health, Uganda (2010). *Implementation Guidelines. Integrated Community Case Management of Childhood Malaria, Pneumonia and Diarrhoea*. Retrieved 29 January 2017, from <https://www.k4health.org/sites/default/files/ICCM%20Implementation%20Guidelines.pdf>
  25. Ngabo, F., Levin, A., Wang, S. A., Gatera, M., Rugambwa, C., Kayonga, C., ... & Hutubessy, R.. A cost comparison of introducing and delivering pneumococcal, rotavirus and human papillomavirus vaccines in Rwanda. *Vaccine*, (2015);33(51), 7357-7363.
  26. Littrell, M., Moukam, L. V., Libite, R., Youmba, J. C., & Baugh, G. (2012). Narrowing the treatment gap with equitable access: mid-term outcomes of a community case management program in Cameroon. *Health policy and planning*, czs110.
- Miller, N. P., Amouzou, A., Tafesse, M., Hazel, E., Legesse, H., Degefie, T., ... & Bryce, J. Integrated community case management of childhood illness in Ethiopia: implementation strength and quality of care. *The American journal of tropical medicine and hygiene*, (2014);13-07