Six Practices to Strengthen Evaluation of Global Health Research for Development

2016

ESSENCE Good practice document series

ESSENCE
on Health Research

Six Practices to Strengthen Evaluation of Global Health Research for Development

2016

ESSENCE Good practice document series

ESSENCE
on Health Research

About ESSENCE and this good practice document

ESSENCE on Health Research is an initiative of funding agencies to improve the coordination and harmonization of research capacity investments. ESSENCE members embrace the principles of donor harmonization and country alignment expressed in the 2005 Paris Declaration on Aid Effectiveness and in the 2008 Accra Agenda for Action. According to these principles, donors strive to align and harmonize their activities and procedures with the priorities of the countries in which they work.

To achieve this goal, ESSENCE members agreed to jointly develop and produce good practice documents that would incorporate current knowledge and best practices on health research and development issues. The first good practice document, called 'Planning, monitoring and evaluation framework for capacity strengthening in health research', was published in 2011 and then subsequently revised in 2016. The second good practice document, called 'Five keys to improving research costing in low- and middle-income countries', was published in 2012. The third good practice document, called "Seven principles for strengthening research capacity in low- and middle-income countries: simple ideas in a complex world", was published in 2014. This fourth document in the series arose from an increasing recognition among ESSENCE members that sharing good practices can help funding organizations, institutes, government agencies and researchers improve efforts to evaluate global health research for development. ESSENCE funders jointly contributed their experiences and good practices in the evaluation of health research for development. ESSENCE consulted researchers, research leaders and other stakeholders at various key points of the development of this document, including at meetings in Ottawa, Canada, Cape Town, South Africa, Windhoek, Namibia, Stockholm, Sweden and Geneva, Switzerland. The aim of this document is to translate all this knowledge into a set of good practices that can help policy-makers, commissioners of evaluations, evaluators, university and government researchers and other actors in overlapping domains of health research and development strengthen their efforts. Although the ESSENCE group is currently health focused, we hope that this document has wider reach and can be used across all research fields.

ESSENCE steering committee

Swedish International Development Cooperation Agency (Sida)
Wellcome Trust
US National Institutions of Health Fogarty International Centre (NIH/FIC)
Canada's International Development Research Centre (IDRC)
UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR)

ESSENCE members

African Development Bank; Association of Commonwealth Universities (ACU); Doris Duke Foundation; European and Developing Countries Clinical Trials Partnership (EDCTP); European Commission (EC); FIOCRUZ, Brazil; German Ministry of Development Co-operation (BMZ/ GIZ); Howard Hughes Medical Institute; Canada's International Development Research Centre (IDRC); Institut Pasteur France; Medical Research Council of South Africa; Ministry of Foreign Affairs of Denmark (Danida); Ministry of Health, Labour and Welfare, Japan; Netherlands Organisation for Scientific Research (NWO/WOTRO); Norwegian Agency for Development Cooperation (Norad); Rockefeller Foundation; Canada's Global Health Research Initiative (GHRI); Swedish Development Co-operation Agency (Sida); Swiss Agency for Development and Cooperation (SDC); the Special Programme for Research and Training in Tropical Diseases (TDR); UBS Optimus Foundation; UK Department for International Development (DFID); UK Collaborative on Development Sciences (UKCDS); US National Institutes of Health/Fogarty International Center (NIH/FIC); US Office of Global AIDS Coordinator; Wellcome Trust; US Agency for International Development (USAID).

For further information on this document and the ESSENCE initiative, please contact:

ESSENCE Chair: Professor Hannah Akuffo, Swedish International Development Cooperation Agency (Sida) at **hannah.akuffo@sida.se**

ESSENCE Secretariat Coordinator: Dr Garry Aslanyan, Special Programme for Research and Training in Tropical Diseases (TDR) hosted at the World Health Organization (WHO) and co-sponsored by the United Nations Children's Fund (UNICEF), the United Nations Development Programme (UNDP), the World Bank and WHO at **aslanyang@who.int**.

Contents

PART I: Evaluating research for development

Introduction

Evaluating research for development, including global health research, is a rapidly changing, complex area of work. This good practice document focuses on Evaluating Global Health Research for Development (GHR4D) that aims to improve equitable health outcomes and sustained well-being in populations around the world through a multidisciplinary, problem-focused approach to research and practice. As GHR4D projects, programmes and initiatives grow in number and complexity worldwide, so does the need for effective evaluation of these endeavours. Recent trends in funding for GHR4D show an increase in the number of funded projects, greater variation in their size and scope and a widening range of funding modalities (Dieleman et al., 2014). At the same time, funders of GHR4D increasingly receive critical questions about the quality and impact of their investment in global health research. Many funders and other stakeholders see evaluation as a suitable tool to help answer these questions. Evaluation can justify continued support, inform research design, improve the management of projects and programmes and contribute to organizational learning (Guinea et al., 2015). In GHR4D, evaluation can also provide an accountability mechanism and serve as an empowerment process. Many available resources explain how to conduct evaluations of health research (Varkevisser C, Pathmanathan I and Brownlee A, 2003) and development (OECD 2010, 2013). But individuals and organizations that commission evaluations express a need for practical guidance which focuses specifically on evaluation in GHR4D. In response, this document presents a set of good practices to inform GHR4D evaluation.

Distinguishing characteristics of GHR4D

Effective evaluation of GHR4D requires an understanding and awareness of characteristics that distinguish GHR4D from other areas. The ways in which researchers conceptualise global health issues and the methodologies that researchers employ, continue to advance and diversify (Gilson, 2012). In addition to conventional disease-based and epidemiological approaches, GHR4D evaluators now complement these strategies with systems approaches, multidisciplinary methods and an emphasis on health equity. This understanding of GHR4D implies attention to strengthening institutions and, more broadly, to systems-level thinking about knowledge production and resource utilisation in national and regional settings. Although some of these characteristics apply to all forms of research for development, they are especially pertinent to GHR4D. Characteristics of GHR4D include:

- **Demand driven:** GHR4D addresses pressing needs in countries where the work takes place. This requires engagement with national and subnational decision-makers, researchers and practitioners, as well as an awareness of and responsiveness to national policies and priorities.

- **Problem-focused, not methods-driven:** The central research question (or problem) that drives a GHR4D project should also inform the approach, design and methods that will be used to evaluate the project. To effectively create a problem-focused evaluation design, one must take into account how social, cultural and political contexts

influence the GHR4D project under study, as well as how these contexts may shape the evaluation process itself.

- **Systems approaches:** In addition to traditional disease-based and epidemiological approaches, GHR4D evaluation benefits from systems-level approaches. Designing a useful and rigorous evaluation for these types of efforts requires appreciation and knowledge of the complexity of systems processes.

- **Multidisciplinary, interdisciplinary, transdisciplinary and rigor:** GHR4D research often draws on multiple disciplinary approaches (e.g. anthropology, biomedicine, economics, epidemiology, political science, public health and sociology). Evaluations of GHR4D research should be open to

and capable of mounting a rigorous assessment of the rationale, methods and outcomes of these complex approaches.

- **Equity:** Many GHR4D projects seek to understand, measure and redress health inequities. Evaluations of GHR4D frequently examine how these projects conceptualise, measure and respond to various axes of social inequity, including disadvantage, deprivation, or vulnerability relating to socioeconomic status, race, ethnicity, gender, sexual orientation, occupation, area of residence, education level and physical or mental disabilities.

- **Ethics:** Ethical health research relies on essential practices such as community participation, informed consent and shared benefits and burdens. These entail a sound understanding of institutional protocols to protect research subjects and issues specific to research in LMICs.

Conditions that create an enabling environment for GHR4D evaluation

Evaluations of global health research for development/GHR4D programmes require strengthened capacities across many actors involved at different stages of the process. These include local evaluators who understand the context and complexity of GHR4D; researchers who understand the purpose and processes involved in participating in an evaluation and who appreciate that external evaluations are not always inherently better than internal

evaluations; commissioners who recognize that it is often advantageous to plan an evaluation early in the process to ensure that evaluators and relevant partners have an opportunity to change the project or programme while there is still time to do so; evaluators who understand that the purpose of the evaluation, cultural appropriateness and considerations of rigor should drive the evaluation design; evaluators who allow the purpose of the evaluation to determine whether the focus will be on outcomes, processes, or both; and GHR4D funders who understand why and how to support evaluations that can prove the effectiveness of an intervention as well as improve similar efforts in the future.

The success of efforts to strengthen capacities among evaluators, researchers and funders also depends on conditions in the countries and locales where evaluations take place (Sombie et al., 2013). Thus, an enabling environment for evaluation in GHR4D also includes national, regional, provincial and local government commitment, local leadership and ownership of the products of health research evaluation and sustained financial support within countries. These country-specific characteristics also facilitate the replication and scale-up of proven interventions and impart constructive lessons learned from attempts that failed or only partly succeeded.

How this good practice document evolved

This document is the outcome of a collaborative process that involved several contributors. In June 2014, IDRC invited several representatives of organizations from the ESSENCE on Health Research Group to meet in Ottawa, Canada. The purpose of this meeting was to identify practices and challenges that relate to research evaluation. In September of 2014, evaluators and other researchers offered their comments on an early draft of the document that ESSENCE members shared during a facilitated discussion at the Third Global Symposium on Health Systems Research in Cape Town, South Africa. Following this meeting, the document was revised to improve its focus, content and usefulness. In early 2015, ESSENCE members contributed additional comments and examples. Finally, in December 2015, further input was solicited during a workshop of the Southern African Research and Innovation Management Association (SARIMA), which took place at the Namibian Institute of Public Administration and Management (NIPAM) in Windhoek, Namibia.

Using this document

The primary intended audience of this document is the organizations that fund and support research for development, including GHR4D and which therefore have a stake in the effective evaluation of GHR4D projects. This may include commissioners of evaluations, who often define what is to be evaluated, actors to be involved, what results are considered to be important, what success and failure look like and what evidence is relevant (IDRC, 2014). Other users may include research evaluators, GHR4D researchers and implementation staff and anyone interested in research assessment, learning and results.

Six practices to strengthen GHR4D evaluation

This document deliberately refrains from the language of 'best practices' and instead presents a set of 'good practices to consider' which is also in line with the other documents produced by ESSENCE. This is meant to acknowledge that questions about what constitutes an 'ideal' evaluation in GHR4D are not necessarily already settled. To present 'best practices' around evaluation in GHR4D would be to oversimplify an area of work that is swiftly emerging as multi-faceted. What works well for one evaluation commissioner may not work for another. Accordingly, this document is framed as a presentation of good practices that pertain to many systemic issues that recur over the lifecycle of a GHR4D evaluation.

The six practices are:

1 | Build on established evaluation standards.

2 | Develop rigorous design, approaches, methods and metrics.

3 | Consider values.

4 | Identify users and intended uses of the evaluation.

5 | Plan according to the anticipated timeline of influence.

6 | Foster collaboration among diverse stakeholders.

PART II : Putting principles into practice

Every GHR4D evaluation happens in its own specific context. With this in mind, the six practices that this document describes are intended to be broadly applicable. One or more examples are provided to demonstrate the significance of each practice as revealed through real-world cases of GHR4D evaluation. Following each example, a selection of additional resources point the way to further information that may assist funders, evaluators and their partners.

Practice 1 | Build on established evaluation standards

 "Failure comes only when we forget our ideals and objectives and principles." Jawaharlal Nehru, 1st Prime Minister of India

Effective evaluation of GHR4D projects is a challenge. Even the most seasoned evaluation commissioners may have doubts about how to proceed in this complex area of research and practice. However, the process becomes easier and yields higher quality results when one keeps in mind a set of time-tested, established evaluation standards and recognizes GHR4D's distinguishing characteristics. The following five established and widely recognized evaluation standards (JCSEE, 1994) can serve as a guide at every stage of the process:

5 standards for evaluation

1. UTILITY | **to increase the extent to which programme stakeholders find evaluation processes and products valuable in meeting their needs.**

2. FEASIBILITY | **to increase evaluation effectiveness and efficiency.**

3. PROPRIETY | **to support what is proper, fair, legal, right and just in evaluations.**

4. ACCURACY | **to increase the dependability and truthfulness of evaluation representations, propositions and findings, especially those that support interpretations and judgments about quality.**

5. EVALUATION ACCOUNTABILITY | **to encourage adequate documentation of evaluations and a meta-evaluative perspective that focuses on improvement and accountability for evaluation processes and products.**

These standards apply throughout the lifecycle of an evaluation. They serve as an important conceptual scaffold for thinking about how to achieve the evaluation's ultimate learning and/or accountability objective. The credibility, accuracy and usefulness of results of an evaluation can improve when one addresses and, if necessary, redresses issues related to these standards sooner in the process rather than later. Evaluation standards also serve to trigger key questions throughout the commissioning process.

These questions include:

- Who is going to use the results and for what purpose?
- When do I need the results? Are there sufficient available resources (time, money and people) to carry out the evaluation?
- Are the planned activities ethical in their design and execution?
- How accurate and dependable are the results?

- Have the relevant processes and milestones been documented for accountability purposes?

Asking these questions often prompts debates that can reveal incongruities between understanding the evaluation's purpose on the one hand and ideas about how to design the research, assumptions about how to implement the evaluation and expectations about how to use the results on the other hand.

Example

United States Agency for International Development (USAID)
Design and modification of a peer review technical panel

The USAID Partnerships for Enhanced Engagement in Research (PEER) Programme[1] supports research capacity building to enable locally-driven solutions to development challenges. Supporting high quality and high impact evidence-based global health research is a critical component of PEER but projects must also have a strong capacity building component and advance USAID's development goals. To address this unique blend of attributes, USAID's US Global Development Lab (known as "Lab")[2] created a unique proposal review process to ensure that projects selected for PEER funding align with the PEER programme's goals. This process also highlights the established standards for evaluation within the GHR4D community.

PEER's unique approach exemplifies the established evaluation standards of utility, feasibility and accuracy. To adequately account for an LMIC context, the PEER programme needed to develop an approach that modified the standard research proposal review process currently in use by other US agencies. The US National Institutes of Health (NIH) and the National Science Foundation (NSF) use 'intellectual merit' and 'broader impacts' to assess applicants' scientific worthiness and potential to make significant advancements in research. In addition to these assessments, PEER's modified proposal review process considers five other areas that correspond to characteristics of GHR4D evaluation. These five additional areas of the PEER review process (with corresponding GHR4D characteristics in bold) are:

1. Need for evidence-to-action in the proposed development challenge (**demand driven**)
2. Development impact of the research activities (**problem focused**)
3. Local stakeholder engagement (**demand driven** and **system approach**)
4. Capacity building (**system approach** and **equity**)
5. Strength of the partnership between the in-country scientists and their US Government funded 'PEER Partner Principal Investigator' (**multidisciplinary** and **rigor**).

To perform the technical review, the implementer of the PEER Programme (the US National Academy of Sciences, Engineering and Medicine) created a two stage qualitative scoring process. In the first stage, the USAID office in an applicant's home country reviews the proposal to determine if the research would be relevant to the country's development priorities (**demand driven** and **problem focused**). These in-country experts comment on technical aspects of the project, intended development impact, feasibility and the project's capacity building potential. In some cases, Operating Units in Washington DC also weigh in. In the second stage of review, an external panel of scientific experts – many with extensive work experience in LMICs – assess the proposal. This panel's review follows the conventional NIH/NSF format, supplemented with an assessment of the US funded partner's strengths and added value to the project (**feasibility standard**). The technical panel's comments result in a final score.

The PEER Programme funds projects that score well in both stages of review. This two-step proposal evaluation method conforms with established standards and responds to GHR4D characteristics. It also increases the likelihood that PEER projects will generate evidence that leads to well-informed policies and builds research capacity in USAID priority countries.

Questions to inform good practice

1. Have you considered the five established evaluation standards?

2. Practically speaking, how do these standards apply to your evaluation?

3. If you make trade-offs applying any standard, how will you mitigate collateral risks?

[1] For information on PEER see: **http://sites.nationalacademies.org/PGA/PEER/index.htm**

[2] For information on US Global Development Lab see: **http://sites.nationalacademies.org/PGA/PEER/index.htm**

Additional resources

In addition to the standards above, the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) provides useful resources:

- A list of 'Criteria for Evaluating Development Assistance'. **http://www.oecd.org/ dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm**

- OECD, DAC (1991) Principles for evaluation of development assistance, Paris: DAC. **http://www.oecd.org/dac/evaluation/dcdndep/41029845.pdf**

- OECD, DAC (1998) Review of the DAC principles of development assistance. Paris: DAC Working Party on Aid Evaluation. **http://www.oecd.org/dac/evaluation/2065863.pdf**

The American Evaluation Association (AEA) also provides guidance on evaluation standards:

- The Program Evaluation Standards. **http://www.eval.org/p/cm/ld/fid=103**
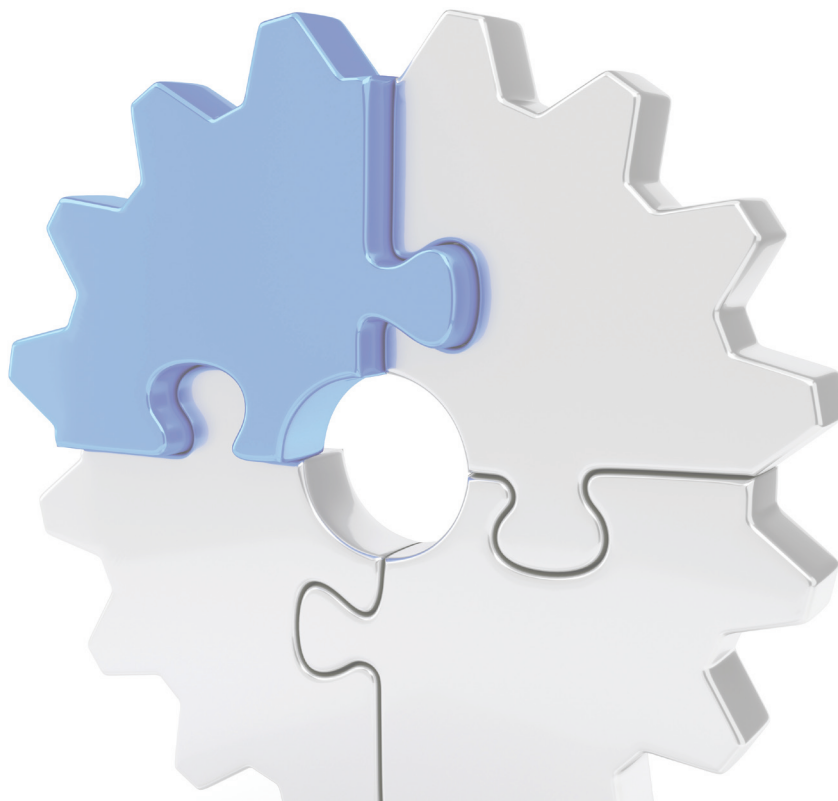
Practice 2 | Develop rigorous design, approaches, methods, & metrics

 *"Research is **formalized** curiosity. It is poking and prying with a **purpose**"*. Zora Neale Hurston, anthropologist and alumna of Columbia University

Successful evaluation requires appropriate approaches, methods and metrics. Yet none of these on their own can drive an evaluation: no single approach, method, or metric (or indicator) is the 'gold standard' for evaluation. Rather, mindful recognition of the users, purpose, context and questions to be asked, as well as adherence to established evaluation standards, should inform an evaluation's design. Depending on the nature of the evaluation, it could adopt an approach rooted in health equity frameworks, biomedical perspectives, a combination of these approaches, or an altogether different frame of reference. Many evaluations use mixed methods, which combine qualitative and quantitative methods to maximize complementary strengths and compensate for limitations. Although mixed methods can bolster the validity and reliability of an evaluation's process and findings, it is important to note that there will likely be trade-offs in rigor, scope, time available to conduct the research and the usefulness of findings.

Example

National Institutes of Health (NIH)

Case study as a method to measure distal impacts of capacity building

The US National Institutes of Health (NIH) incorporates the use of case-studies in its methodological approaches to studying capacity building, which is an important part of broader efforts to improve global health. Quantitative methods such as counting numbers of publications and successful grant applications are among the most commonly used metrics to evaluate Research Capacity Strengthening (RCS). Yet thorough and accurate assessment of capacity building, including RCS impact on health systems, remains an imprecise enterprise.

The Fogarty International Center (FIC) at the NIH utilizes case study methodology to measure distal (longer-term, or "farther down the road") impacts of capacity building[3]. This approach provides a more comprehensive understanding of the impact of investments on in country networks, research training, institutional culture of science and the ways in which investments produce evidence to inform policy and programmes. Case studies utilized telephone interviews to document practical applications of research findings, periodic progress reports and survey responses from individuals. Although time intensive, the case studies provide important qualitative information about the impact of investments in health research. Several of the case studies tell success stories that quantitative measures alone could not measure or communicate.

[3] To learn more about how NIH/FIC used case-study methodology, see: **http://www.fic.nih.gov/About/Staff/Policy-Planning-Evaluation/Documents/fogarty-global-brain-disorders-program-review-2014.pdf**

Example

South African Medical Research Council (SAMRC)
Mixed methods to determine both implementation success and financial sustainability

SAMRC took a mixed methods approach to evaluate the effect of the UNICEF Integrated Health Systems Strengthening programme (IHSS) – which aimed to increase coverage of low cost maternal and child health interventions – on 'The Catalytic Initiative to Save a Million Lives' (CI), an international partnership to accelerate progress on the health-related Millennium Development Goals (MDGs) in six African countries. SAMRC's evaluation[4] aimed to determine costs of implementing Integrated Community Case Management of Common Childhood Illnesses (iCCM) and learn if iCCM scale-up would be sustainable after IHSS support ended.

The evaluation's combination of quantitative and qualitative methods allowed it to measure socioeconomic progress, epidemiological changes and effects of other donor and government interventions. Collectively, these indicators allowed an investigation of the extent to which changes in child mortality could be attributed to increases in iCCM coverage. Analysis suggested that significant reductions in the Under Five Mortality Rate (U5MR) could be achieved by strengthening case management. Yet interview data revealed concerns about the affordability of sustaining the programme: the report quoted a Ministry of Health official in Niger as saying "Free care is very popular, but this government is struggling to continue." However, in March 2014 UNICEF, the Global Fund to Fight AIDS, Tuberculosis and Malaria and the Reproductive, Maternal, Newborn & Child Health (RMNCH) Trust Fund announced plans to scale-up iCCM. The evaluation thus concluded that Catalytic Initiative indeed had fulfilled its 'catalytic' role.

---

[4] Doherty et al. (2014) Report on Summative External Evaluation of the Catalytic Initiative (CI)/ Integrated Health Systems Strengthening (IHSS) Programme **http://www.mrc.ac.za/healthsystems/ IHSSMultiCountryReport.pdf**

Questions to inform good practice

1. Will the evaluation design allow you to answer your evaluation questions?

2. Do the approach, methods and metrics align with the needs of the users? If so, are they feasible and will they yield results that are easy to understand and put into practice?

3. What are the ethical implications of the evaluation design? How will these be addressed?

Additional resources

- For a knowledge-sharing platform that organizes design, methodological and metrics-related evaluation tasks into user-friendly clusters see:
   http://www.betterevaluation.org

- Cole, Donald et al. (2014) Indicators for tracking programmes to strengthen health research capacity in lower- and middle-income countries: a qualitative synthesis, Health Research Policy and Systems, 12:17.
   http://www.health-policy-systems.com/content/12/1/17

- Grazier, Kyle L et al. (2013) Estimating return on investment in translational research: Methods and protocols. Evaluation & the Health Professions, 36(4): 478-491.
   http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4084908/

Practice 3 | Consider values

"We are, at almost every point of our day, immersed in cultural diversity: faces, clothes, smells, attitudes, values, traditions, behaviours, beliefs, rituals". Randa Abdel-Fattah, writer, lawyer and champion for social justice

The word 'value' is at the root of evaluation and for good reason: not only can the values that are associated with political contexts influence evaluation, but the values of individuals and groups that commission and conduct evaluations can influence political situations. GHR4D evaluation design requires careful consideration of power relations among constituencies, critics, commissioners and other influential groups. It is important for GHR4D evaluation to recognize, measure and respond to the various values that shape this political context – a context which comprises economic, social, cultural, technical, legal and ethical factors.

Reflection on ethical factors helps illustrate the role that values play within GHR4D's broader political context. Consider, for example, the role of reciprocal influence among various groups, each of whose individual actors have their own personal and professional values. While reciprocal influence can strengthen intersectoral bonds and lead to greater accountability for ethical practice, other types of influence can open the possibility that some groups and individuals will suffer consequences that would be unethical or inconsistent with the 'Propriety' standard (see Practice 1, "Five standards for evaluation"). An honest reckoning with the full range of values that may influence GHR4D evaluation requires identifying the primary intended user(s) and taking an inventory of the core values that are intended to drive the evaluation.

Identifying the groups whose values will influence the evaluation – and understanding the extent and implications of their influence - can help answer key questions: Whose values are being measured? How and with whom will these values be negotiated? How will different groups value positive and negative outcomes of the GHR4D project? Keep in mind that the organizational values of the commissioner's own organization also influence the design, interpretation and use of evaluation findings. Moreover, values and approaches to measuring them vary significantly among evaluation commissioners. For example, while one evaluation may value return on investment – a '4E' framework (economy, efficiency, effectiveness, equity) is often used to evaluate 'VfM' (value for money) - and therefore prioritize the measurement of satisfactory outputs produced at the lowest possible cost (e.g. number of midwives trained), another evaluation of the same initiative may place more value on measuring equitable distribution of benefits (e.g. midwives' use of training to reach vulnerable populations).

There is no right or wrong way to define or measure values. What matters is to discuss values early on, revisit these issues periodically and strive for clarity among intended users. This helps an evaluation pursue a suitable engagement strategy and translate findings in an effective and sensitive manner, but also strengthens the evaluation's ability to uphold principles of ethical research, including community participation, informed consent and shared benefits.

Example

International Development Research Centre (IDRC)
Developmental evaluation for cross-grant learning and exchange

In 2014, Canada's IDRC funded a group of seven research grants under the SEARCH project (Strengthening Equity through Applied Research Capacity Building in e-health). These grants support implementation research projects in Bangladesh, Burkina Faso, Ethiopia, Kenya, Lebanon, Peru and Viet Nam. IDRC designed an evaluation of these projects to examine issues of health equity, governance and systems integration. Crucially, this evaluation aimed to explicitly acknowledge and respond to values and the measurement of values.

The evaluation adopted a developmental process approach, which is a strategy that grants equal importance to unexpected, negative and positive outcomes[5]. This approach aligned well with the values of the funding organization, as well as with the novel and exploratory nature of the research projects. It allowed users of the evaluation to continuously inform the trajectory, analysis and use of evaluation findings as they emerged. To accomplish this, the evaluators created critical points of reflection, periodically provided feedback to SEARCH Project staff and collaborated closely with the project staff to facilitate appropriate action. One finding that emerged from this was a need among project staff to strengthen their capacities for gender analysis. Subsequent to this finding, IDRC programme officers identified suitable support strategies and mechanisms to respond to this need. Thus the evaluation scanned for values on what otherwise might have remained 'unexplored avenues'.

---

[5] For more information on evaluation at IDRC, see: **http://www.idrc.ca/EN/Documents/Evaluation-at-IDRC.pdf** and **http://idl-bnc.idrc.ca/dspace/bitstream/10625/47288/1/133634.pdf**

Questions to inform good practice

1. What values steer the evaluation and whose values are they?

2. How does the evaluation respond to or measure values?

3. What is the impact(s) on different groups of the evaluation process and findings?

4. Does the evaluation carefully consider and address ethical concerns?

Additional resources

-   Morris, Michael (Ed.) (2007) Evaluation Ethics for Best Practice: Cases and Commentaries. New York: Guilford Press.

-   Östlin, Piroska et al. (2011) Priorities of Research on Equity and Health: Towards an Equity-Focused Health Research Agenda, PLoS Medicine, 8(11): e1001115 **http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001115**

-   Weiss, Carol Hirschon (1991) Evaluation research in the political context: Sixteen years and four administrations later, Evaluation and Education: At Quarter Century, McLaughlin, Milbrey W and Phillips, D (Eds.) Chicago: University of Chicago Press.

-   House, Ernest R. (2011) Evaluating with Validity. Information Age Publishing.

-   House, Ernest R, Howe, Kenneth R. (1999) Values in Evaluation and Social Research. Sage Publications.

Practice 4 | Identify users and intended uses of the evaluation

"A gift consists not in what is done or given, but in the intention of the giver or doer". Seneca, Moral Essays, Volume III: de Beneficiis

There are numerous ways to approach the process of commissioning an evaluation for GHR4D. Although many of these approaches overlap, their differences hold the potential to influence key decisions. It is helpful to picture evaluation as a tree with three distinct but entwined branches. One branch prioritizes methods (see Practice 2). Another concerns values (see Practice 3). The third branch grows around use. While it is important to consider all three branches of any evaluation design, it is essential to recognize that each of these branches is more strongly rooted in some of the five established standards for evaluation than others. It follows that weighting any one branch more than the others is likely to alter the 'evaluation tree's' overall trajectory of growth. Consider, for example, the use branch. It corresponds most closely to the standard of 'utility'. A focus on utility should not preclude attention to the other four standards of evaluation (i.e. feasibility, propriety, accuracy and accountability), yet the practice of identifying users and intended uses of an evaluation lays the groundwork for key decisions about whether and how to take actions or to implement reforms in GHR4D projects.

Identifying concrete users for evaluation findings and understanding how these users would likely use the findings, should be established early on and then revisited over the course of the evaluation. As no evaluation is 'value-free', the process of identifying users and uses can elucidate whose values will frame the evaluation and who will have the responsibility to apply findings and implement recommendations (Patton, 2008). For example, evaluating a particular GHR4D project will take on a different trajectory if the primary intended user is: (a) interested in using the findings to publish academic papers and influence academic discourse; or (b) a local community leader who is interested in changing practices related to health service delivery.

In collaboration with other stakeholders, a GHR4D evaluation commissioner (who may also represent one of the user groups) will need to decide how to identify primary intended user(s), how to involve them in the evaluation process and how to ensure their uses inform significant decision points. This process in itself is very useful for everyone involved to gain a shared understanding regarding the purpose of the evaluation.

With the primary intended users and their intended uses serving as an anchor for the evaluation, there should be sufficient flexibility built in to operational aspects of the study to allow processes and activities to adapt to changes over the course of the evaluation. Flexibility entails negotiation between the evaluation commissioner, the primary intended users, the evaluators and the subjects of the evaluation.

Example

Swedish International Development Cooperation Agency (SIDA)

Users and uses of a research capacity strengthening (RCS) programme

In 2009, Sida decided to initiate an evaluation of its university-wide support for the Research Capacity Strengthening (RCS) programme at Makerere University in Uganda[6]. This was a multi-faceted evaluation study of ongoing efforts to support strengthening of the University's institutional research capacities. As the primary intended user of the evaluation, Sida had already set up a monitoring team in 2004 to develop and institutionalize continuous self-assessment focused on strengthening research capacities, among other tasks. It was the intention from the outset that this monitoring team would also perform an evaluation for Sida. As the primary intended user, Sida was closely involved in developing the evaluation design, instruments and analytical tools. The decision to engage evaluators from the existing monitoring team proved to be strategic and effective, as these individuals were well-informed of the context and complexities of both the institution and the programme of work.

Sida used findings from the evaluation to decide if and how to continue supporting processes to improve research capacities within Makerere University. Based on these findings, which demonstrated improvements in several areas – most notably in strengthening the culture of research – Sida decided to renew its funding support for 2010-2014. As a secondary user of the evaluation, Makerere University used the findings to initiate institutional changes, including addressing capacity gaps in the University's administration, financial management and programme governance. The University's subsequent research and reforms ultimately also attracted other funding (from IDRC) to implement reforms.

---

[6] Enhancing Research Capacity at Makerere University, Uganda through collaboration with Swedish Universities, 2000–2008, Sida, 2010 **http://www.sida.se/contentassets/261ae901c4824efb98468e4186ff5234/ enhancing-research-capacity-at-makerere-university-uganda-through-collaboration-with-swedish-universities-2000-2008past-expe_2854.pdf**

Questions to inform good practice

1. Who are the primary users of the evaluation? Who are the secondary users? Who else is in the audience and what is their interest in the evaluation?

2. How are the intended users expected to use the findings?

3. Do the evaluation questions align with the needs of the primary users?

4. Will the data collection and analysis plans match the needs of evaluation users?

Additional resources

- Alkin, Marvin C and Christie, Christina A (2004). An Evaluation Theory Tree, Evaluation Roots: Tracing Theorist's Views and Influences, Alkin, Marvin C (Ed.), Thousand Oaks, CA: Sage. Pp. 12-65. **http://www.sagepub.com/sites/defaultfiles/ upm-binaries/5074_Alkin_Chapter_2.pdf**

- Hallam, A and Bonino, F (2013) Using evaluation for a change: Insights from humanitarian practitioners, ALNAP Study, London: ALNAP/ODI

- Mayne, J (2014) Issues in enhancing evaluation use, Enhancing Evaluation Use: Insights from Internal Evaluation Units, Laubli-Loud, M and Mayne, J (Eds.), Thousand Oaks: Sage

Practice 5 | Plan according to the anticipated timeline of influence

"Nine-tenths of wisdom consists in being wise in time". Theodore Roosevelt: author, naturalist and 26th President of the United States

Close attention to timelines and timeliness is vital to the process of commissioning an evaluation. GHR4D programmes and projects aim to bring about change. Change, however, usually takes time to occur and even more time to measure. Therefore, the process of defining clear expectations for a timeline of influence should begin early in the evaluation commissioning process (Sridharan & Nakaima, 2011). Evaluations often comprise multi-step processes that require time to implement and effort to ensure that decisions align with the overall purpose of the study. Early planning and reflection on the anticipated trajectory of outcomes can assist in the timely input of evaluation results in decision processes.

Timeline-related decisions include identification and anticipation of key milestones and contemplation of the chronology of planned activities. Perhaps the most important decision that relates to the anticipated timeline of influence is when to begin an evaluation. The idea that the design of evaluations should begin only near the end of projects or programmes is a misconception: on the contrary, evaluations can begin at any time; moreover, it is often best to initiate them from the very start of a project. Although there is no rule about when to initiate an evaluation, it is helpful to ensure there is sufficient time to: identify and engage with primary intended users and to learn about their intended uses, understand and respond to values and contexts that shape the particular GHR4D project and establish processes to recognize and respond to opportunities for influence through the sharing of evaluation findings.

An evaluation plan's engagement with the timeline of influence also poses implications for the evaluation's rigor and utility. There is often a time-lag between the anticipated timeline of influence and the actual trajectory of outcomes as it occurs in the real world. For example, although a five-year GHR4D project may have an anticipated timeline of influence that would coincide with Year 4 or Year 5, the actual outcomes may not be realized until after the end of the project or in the subsequent funding cycle. Staying mindful of this reality can help position an evaluation to generate and translate findings in a timely way that achieves maximum influence and impact. Moreover, early planning and reflection on the anticipated trajectory of outcomes can forge links between results and longer term use within and outside of the immediate GHR4D context.

Example

UK Department for International Development (DFID)
Planning an evaluation for the Humanitarian Innovation and Evidence Programme (HIEP)

HIEP is a long-term programme in DFID that allocates funding for projects which seek to generate or synthesize evidence on what works to successfully redress problems affecting humanitarian effectiveness. HIEP includes projects in key areas such as health in emergencies, disaster risk reduction, scaling up cash-based responses, working in volatile environments and urban resilience.

A formative evaluation[7] of HIEP was conducted between January and May 2014 to check the feasibility and appropriateness of eight of the 20 HIEP-funded projects. This formative evaluation made an initial assessment of HIEP design, identified progress to date and made recommendations to facilitate learning. The formative evaluation also provided a foundation for two phases of summative evaluation, during which DFID intends to share draft case study reports with its teams and partners for fact-checking as a way to ensure findings and recommendations reach them more quickly.

The timetable of the evaluation was developed to meet the needs of DFID as the principal user. Rather than wait until the midpoint or end of this programme, DFID's evaluation of HIEP was designed to be an ongoing four phase, five year (2013-2018) evaluation that commenced from the outset. This design provides several benefits. First, it allowed the evaluation team to begin their work early in the process. Second, it provided flexibility for the evaluation team to respond to changes in activities and context in the field as they arose, while also allowing the team to adhere to a relatively firm timeline for the different phases of the study. Third, it placed the evaluation team in a position to articulate and test a fuller programme Theory of Change (or a comprehensive description of how and why a desired change is expected to happen in a particular context). The Theory of Change that emerged subsequently served as a critical 'touchstone' to help guide the evaluation and interpret its process and findings as they unfold over time.

---

[7] Evaluation of the Humanitarian Innovation and Evidence Programme (HIEP): Formative Phase Report. DFID (Department for International Development). Itad, 2014. **http://r4d.dfid.gov.uk/Search/ DocumentLinks.aspx?OutputID=200759**

Example

Norwegian Agency for Development Cooperation (NORAD)

Early evaluation to catch and redress omissions that are fundamental to a programme

The Health Results Innovation Trust Fund (HRITF) is a multi-donor Trust Fund managed by the World Bank with financial support from the Government of Norway and, from 2009, the UK's Department for International Development (DFID). The Fund's purpose is to support pilot projects that use Results Based Financing (RBF) approaches in the health sector, including the quantity and quality of reproductive, maternal, neonatal and child health. RBF – also known as "performance based incentives/PBI" or "pay for performance/P4P" – is a funding scheme that links incentives with measurable results.

NORAD's Evaluation Department commissioned an evaluation[8] that focused on the first four years of HRITF (2007 to 2011). This evaluation, which was the first of three that will take place over the course of the Fund (2007 to 2020), assessed HRITF's performance against its objectives and made recommendations to improve its operations, programming and governance. The evaluation process began in the early years of HRITF. This timing enabled the identification of areas which required change but which could easily be addressed soon enough to make a critical difference. One of the evaluation's most significant discoveries was that the Fund did not have sufficient indicators to measure success. This was a crucial and timely finding, given that HRITF operates on the RBF model: an explicit premise for the Fund's success is the usage of indicators to measure progress and reward activities that move the Fund closer to its objectives. While the evaluation also documented many activities that are encouraging, it was essential to ask whether these were the right activities and whether the manner of their implementation would move HRITF closer to its goals. The evaluation's anticipation of HRITF's timeline of influence allowed it to detect this fundamental limitation in the Fund's operational design early in the process. This allowed HRITF actors to develop a framework that explicitly defines results the Fund expects to achieve to fulfil its overarching objectives and that includes programme level indicators to measure such progress.

---

8   Martinez J, Pearson M, Sørensen BH, James B, Sambo C (2012) Evaluation of the Health Results Innovation Trust Fund. NORAD (Norwegian Agency for Development Cooperation), Department of Evaluation. Report 4 (June) **http://www.norad.no/globalassets/import-2162015-80434-am/www.norad.no-ny/filarkiv/vedlegg-til-publikasjoner/hritf_lr3.pdf**

Questions to inform good practice

1. Has the process of commissioning an evaluation started early enough to identify and respond to the evaluation users and their intended uses? If not, how can timelines be adjusted to maximize impact of the evaluation?

2. What is a realistic anticipated timeline of influence for the intervention being evaluated and the evaluation findings to be generated?

3. What strategies and tools can be used to alter the schedule of evaluation activities so that the findings may be produced and shared in a timely manner?

Additional resources

- The Center for Theory of Change provides an overview of concepts and resources for implementation of a Theory of Change approach. **http://www.theoryofchange.org**

- Sridharan, Sanjeev et al. (2006) Developing a stakeholder-driven timeline of change for evaluations of social programs. American Journal of Evaluation, 27(2): 148–162 **http://aje.sagepub.com/content/27/2/148.abstract**

Practice 6 | Foster collaboration among diverse stakeholders

"If you want to go fast, go alone. If you want to go far, go together".
African proverb

Evaluation in GHR4D benefits from skilful navigation of interpersonal and inter-organizational politics and dynamics. This can refer to the interactions among commissioners (if the project is financed by multiple funders), as well as the interactions between the commissioner(s) and the evaluator. If one were to draw a continuum of the intensity of these interactions, it could have the following two extremes: at one end, a high level of independence (i.e. low level of engagement) between the evaluator and the evaluation commissioner(s); at the other end, a high level of engagement between the evaluator and evaluation commissioner(s) (see Figure 1).

High levels of engagement among commissioners, evaluators and other stakeholders in GHR4D evaluations are beneficial when they create a positive effect. However, to achieve a positive effect all parties should possess a shared understanding of their respective roles and of the evaluation's purpose. Yet the roles that different groups of individuals play and the central purpose of the evaluation tend to vary: it is normal for them change over the lifetime of the evaluation, but this variability may need to be periodically revisited over the course of the evaluation process. The uncertainties that accompany this fluid and dynamic process can become opportunities to engender trust and goodwill among diverse stakeholders if collaboration and communication is strong.

Factors to consider when deciding where to place an evaluation on this continuum and how much engagement and interaction should take place among stakeholders include:

- Ensuring that the level of engagement aligns with established evaluation standards.
- Recognizing and responding to values and the context that shapes the evaluation process.
- Keeping in mind the users, uses and purpose of the evaluation.
- Being mindful of the time and resources that are available for the evaluation and how to best allocate these in the commissioner-evaluator engagement.
- Balancing involvement of various stakeholders to maximize timely and sustained influence.
- Checking that the level of engagement will strengthen the proposed evaluation design, methods and subsequent analyses.

Figure 1: Continuum of engagement between evaluation commissioners and evaluators

Low level
of engagement

High level
of engagement
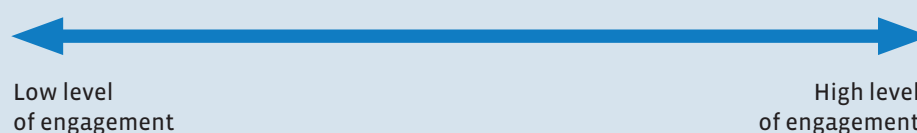
These factors demonstrate the interdependence and mutual reinforcement of the six practices that this document advocates. Awareness of these linkages among good practices can improve the design and practice of GHR4D evaluation in real-world situations.

In cases where there is more than one commissioner (e.g. if a project has received support from multiple funders), it is helpful to identify the values, objectives, roles, ways of working and expectations of all the different commissioners. This requires strong collaboration and communication skills among the commissioners. These different types of alignments and interactions can be visualized on a different sort of continuum (see Figure 2), one with 'networking' at one extreme and 'partnering' at the other.

Depending on the nature of a multi-funder GHR4D initiative, it may be necessary to form a governance structure for specific functions to allow for strong communication among the various funding agencies, as well as a decision-making mechanism that can facilitate a productive collaboration. Fostering collaboration and networking among the various actors involved in GHR4D evaluation is widely acknowledged to be necessary. The challenge is to effectively apply this practice in the context of the myriad relationships, expectations and processes that are involved. Keys to rising to this challenge include early planning, continual practice, careful attention to the multiple 'moving parts' of evaluation, teamwork and persistent focus on the end-goal of delivering valid and reliable findings.

Figure 2: Spectrum of alignment and interaction among commissioners

A   B   C   D   E
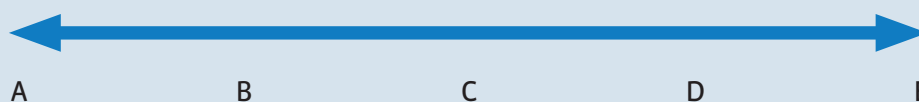
A. Networking (sharing information and ideas)

B. Cooperating (helping distinct members accomplish their separate individual goals)

C. Coordinating (working separately on shared goals)

D. Collaborating (working together towards a common goal but maintaining separate resources and responsibilities)

E. Partnering (shared goals, shared decisions, shared resources within a single entity).

Example

The Newton Fund
Collaborative and networking approach to commissioning an evaluation

The Newton Fund is a £375m programme which was launched by the UK Government in 2014 to develop science and innovation partnerships that promote the economic development and welfare of LMICs. For the United Kingdom it is a different approach to funding science for development: resources are only available through collaboration with partnering government funders; this helps ensures that local challenges and needs are addressed. The Fund works in 15 partner countries to support a wide range of activities –from fellowships, to joint research programmes on health and development challenges, to innovation skills training. Match funding with partnering countries and national priority setting have generated a plethora of science and innovation activities for development models in GHR4D and beyond.

Given the scope and complexity of this initiative, the Department for Business Innovation and Skills (BIS), which oversees the Fund, was keen to design an evaluation that would integrate effective approaches and practices from other funding organizations and to embed the evaluation as early as possible. There was also awareness that wide consultation would help secure buy-in and ultimately use of the evaluation results.

A working group on evaluation convened senior representatives from UK stakeholders of the Fund. A consultative networking exercise was initiated by UKCDS (the UK Collaborative on Development Sciences) in the summer of 2014. This drew on experts in USAID, the Bill and Melinda Gates Foundation, Canada's International Development Research Centre (IDRC) and others, as well as academic and grey literature. Based on the results, UKCDS authored a paper that provided advice on how to design an evaluation. The plurality of experiences was particularly important given the Newton Fund's relatively new way of funding a complex and multi-faceted programme of work.

In parallel, the Fund's managers led work to harmonise monitoring data collected among the various implementation partners. Evolving from the working group, an expert group consisting of independent evaluation specialists, including the [UK] Department for International Development (DFID), the National Audit Office and academia, was set up to recommend formal processes and to advise on commissioning an overarching impact evaluation. The call for proposals was also peer-reviewed by experts through DFID's quality assurance process. The call for proposals closed in July 2015 and the submitted applications were evaluated by a panel from the expert group. It is anticipated that the networking and collaboration approach with others will minimise 'reinvention of the wheel' and ensure the Newton Fund's evaluation is robust and transparent.

Example

INDEPTH Network

Inter-agency collaboration to assess LMIC research network and inform its strategic plan

In accordance with the Swedish International Development Cooperation Agency's (Sida) policy to regularly evaluate the organizations that it supports, Sida began to plan an evaluation of its investments in 2009. To avoid duplication of efforts – and as an example of collaboration with diverse stakeholders – Sida consulted with other funders about the terms of reference (TORs) for this evaluation to ensure that the results could be used by all the supporting funding agencies.

Of note, the evaluation studied all activities of INDEPTH, which is a network of research centres led by actors in the Global South that conducts health and demographic surveys of geographically-defined populations in low- and middle-income countries. Several funding agencies, including Sida, support this network. This external review was intended to help assess the network's performance and to inform the creation of its next strategic plan. The evaluators shared their report with the INDEPTH secretariat as well as with all of the funding agencies that participated in the network.

Questions to inform good practice

1. Is there a shared understanding of the aims and process of the evaluation between the evaluation commissioner and the evaluator?

2. If multiple commissioners are involved, is there a shared understanding of the aims and processes of the evaluation among the different parties?

3. Is the timeline and budget realistic to allow for joint decision-making processes among the different parties involved in commissioning an evaluation?

Additional resources

- Adam T, Røttingen J and Kieny MP (2015) Informing the establishment of the WHO Global Observatory on Health Research and Development: A call for papers, Health Research Policy and Systems, 13:9

- ICDK (2015) Evaluation of the Danish Innovation Centres. **http://ufm.dk/en/publications/2015/filer-2015/15-022371-02-bilag-3-evaluation-of-the-danish-innovation-centres-9265391_1_1.pdf**

- Patton, Michael Quinn (2008). Utilization-focused Evaluation (4th ed.). Thousand Oaks, CA: Sage

- World Health Organization Maximizing Positive Synergies Collaborative Group (2009) An assessment of interactions between global health initiatives and country health systems, Lancet, 373: 2137-2169 **http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(09)60919-3/fulltext**

PART III : Practical ways forward

The emergence of GHR4D evaluation as means to prove or improve the design and implementation of projects and programmes prompted ESSENCE members to collaboratively formulate and refine the aforementioned practices. This document presents these practices in a manner that deliberately moves away from the instructional approach that characterizes other "best practice" guidebooks and manuals. Instead, the approach here is to recommend that commissioners of evaluations, evaluators and other groups give due consideration to six 'good practices', each of which were thoughtfully selected for inclusion in this document. To ground these good practices in real world cases, this document also presents several examples of the actual processes and applications of GHR4D evaluation. As GHR4D evaluation continues to evolve, it may be useful to think of the ways in which the practices presented here intersect with more than one example.

GHR4D evaluation continues to evolve as a vital area of work that is necessary for evidence-based justifications and targeted improvements of both funding and conduct of research. One key to unlocking the full potential of vitality of work in this area is reliance on the dynamic, interconnected and flexible yet collective and principled approach that the suggested good practices and examples in this document describe.

We hope this document has the potential to:

- Provide guidance to commissioners of GHR4D evaluations, as well as evaluators and other stakeholders in GHR4D projects, programmes and initiatives
- Improve the quality, credibility and usefulness of GHR4D evaluations
- Assist funders of GHR4D efforts by providing information about a tool (i.e., evaluation) that is widely used to judge whether investments 'pay off'.

In conclusion, this document does not intend to represent the 'final word' on evaluation in GHR4D. Rather, its main contribution is the presentation and discussion of selected good practices and the principles that justify them. Of equal importance, it is hoped that this document will serve as a common point of reference for funders, recipients of funding, evaluators and other interested groups to continue the discussion and encourage innovative thinking that can strengthen GHR4D evaluation design and implementation.

Sources

AEA Evaluation Principles: http://www.
eval.org/p/cm/ld/fid=51

Bill and Melinda Gates Foundation
(2014) Methods for Economic Evaluation
Project (MEEP). http://www.nice.org.uk/
about/what-we-do/nice-international/
nice-international-projects/methods-for-
economic-evaluation-project-and-the-
gates-reference-case

Dieleman, Joseph L et al. (2014) "Global
health development assistance remained
steady in 2013 but did not align with
recipients' disease burden", Health Affairs,
33(5):878–886. http://content.healthaffairs.
org/content/33/5/878.full accessed on
08/01/2016

Gilson, Lucy (Ed.) (2012) Health policy
and systems research: a methodology
reader. Alliance for Health Policy and
Systems Research, Geneva: World Health
Organization. http://www.who.int/
alliance-hpsr/alliancehpsr_reader.pdf
accessed on 08/01/2016

Guinea, Joaquin et al (2015) "Impact
oriented monitoring: A new methodology
for monitoring and evaluation of
international public health research
projects", Research Evaluation, 24(2):131-
145. http://rev.oxfordjournals.org/
content/24/2/131 accessed on 08/01/2016

Guthrie S et al. (2013) Measuring Research:
A guide to research evaluation and tools.
RAND Europe. Retrieved from: http://www.
rand.org/pubs/monographs/MG1217.html

IDRC (2014) Guide for Commissioning
Evaluations. http://www.idrc.ca/
EN/Programmes/Evaluation/Pages/
ProjectDetails.aspx?ProjectNumber=10714 9
accessed on 08/01/2016

JCSEE (1994) The Programme Evaluation
Standards, Joint Committee on Standards
for Educational Evaluation. Thousand
Oaks, CA: Sage. https://us.sagepub.com/
en-us/nam/the-programme-evaluation-
standards/book230597 accessed 08/01/2016

OECD, DAC (1991) Criteria for Evaluating
Development Assistance. DAC Principles
for the Evaluation of Development Assis-
tance. http://www.oecd.org/dac/
evaluation/daccriteriaforevaluatingdevel
opmentassistance.htm

OECD (2010) Evaluation in Development
Agencies. http://www.oecd-ilibrary.org/
development/evaluation-in-development-
agencies_9789264094857-en accessed on
08/01/2016

OECD (2013) Evaluating Development
Activities: Twelve Lessons from DAC Peer
Reviews. Organisation for Economic Co-
operation and Development. http://www.
oecd-ilibrary.org/development/evaluating-
development-activities_9789264196360-en
accessed on 08/01/2016

Patton, Michael Quinn (2002) Utilization-focused Evaluation Checklist. http://tamarackcommunity.ca/downloads/vc/ufe_checklist.pdf

Patton, Michel Quinn (2008) Utilization-focused Evaluation (4th ed.), Thousand Oaks, CA: Sage

Sombie, Issiaka et al. (2013) The state of the research for health environment in the ministries of health of the Economic Community of the West African States (ECOWAS). Health Research Policy and Systems, 11: 35. http://www.health-policy-systems.com/content/pdf/1478-4505-11-35.pdf accessed 08//01/2016

Sridharan, Sanjeev and Nakaima, April (2011) Ten steps to making evaluation matter, Evaluation and Programme Planning, 34(2): 135-146.

United Nations Development Programme (UNDP). (2009) Handbook on Planning, Monitoring and Evaluating for Development Results. UNDP. Retrieved from http://web.undp.org/evaluation/evaluations/handbook/english/documents/pme-handbook.pdf

Varkevisser C, Pathmanathan I and Brownlee A (2003), Designing and conducting health systems research projects. Volume ll: Data analysis and report writing. Jointly published by KIT Publishers and the International Development Research Centre (IDRC), in association with the Africa Regional Office (AFRO) of the World Health Organization. http://archives.who.int/prduc2004/Resource_Mats/Designing_2.pdf accessed on 08/01/2016

Annex A:
Additional resources for impact evaluation

Gertler, Paul J et al. (2011) Impact Evaluation in Practice, Washington, DC: The World Bank http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf

Impact evaluation: Overview http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,menuPK:384339~pagePK:162100~piPK:159310~theSitePK:384329,00.html

Kok, Maarten O and Schuit, Albertine J (2012) Contribution mapping: a method for mapping the contribution of research to enhance its impact, Health Research Policy and Systems, 10:21 http://www.health-policy-systems.com/content/10/1/21

Kuruvilla, Shyama et al. (2006) Describing the impact of health research: A Research Impact Framework. BMC Health Services Research, 6:134 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1635046/pdf/1472-6963-6-134.pdf

Luoma, Päivi et al. (2011) Better Results, More Value: A framework for analysing the societal impact of Research and Innovation, Tekes Review 288/2011, Helsinki: Tekes http://www.tekes.fi/globalassets/julkaisut/better_results_more_value.pdf

Ramalingam, Ben (2011) Learning how to learn: eight lessons for impact evaluations that make a difference, ODI Background Note, London: Overseas Development Institute (ODI) http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7096.pdf

Rogers, Patricia J (2012) Introduction to impact evaluation, Impact Evaluation Notes, Number 1 (March), RMIT University (Australia) and Better Evaluation https://www.interaction.org/sites/default/files/1%20-%20Introduction%20to%20Impact%20Evaluation.pdf

UKCDS (2013) Why and what? Motivations for evaluating impact and tools for framing right questions. London http://www.ukcds.org.uk/sites/default/files/content/resources/Evaluating%20the%20Impact%20of%20Research%20Programmes%20-%20Motivations.pdf

ESSENCE
on Health Research

Secretariat is hosted at:

TDR
For research on
diseases of poverty
UNICEF • UNDP • World Bank • WHO