



Perspectives on Evaluating Global Health Research for Development: A Background Paper

Taking stock of current practice and ways to improve it

CANADIAN COALITION FOR GLOBAL HEALTH RESEARCH

Promoting better and more equitable health worldwide through the production and use of knowledge

Vic Neufeld, Donald C. Cole, Alan Boyd, Donald Njelesani, Imelda Bates and Stephen R. Hanney

27 June 2014

Acknowledgements

This background paper was commissioned by Canada's Global Health Research Initiative in collaboration with ESSENCE on Health Research.

The CCGHR team would like to express our sincere thanks to the participants of the GHRI-ESSENCE workshop on Perspectives on Evaluating Global Health Research for Development, held in Ottawa on June 2 & 3, 2014, for their insightful and thoughtful comments on an earlier version of this document.

Contents

Executive Summary.....	i
Introduction.....	1
The Global Health Research for Development (GHR4D) context.....	2
Research Design.....	4
Findings.....	5
A. Current Practices:.....	5
B. Approaches and methods for evaluating research and increasing the use of evaluation results:.....	9
C. Similarities and Differences:.....	13
D. Challenges:.....	15
E. Other findings:.....	18
Discussion.....	19
Selected References.....	23
Attachments.....	24
1. Definitions.....	25
2. Participating Agencies and Interviews Completed.....	26
3. Research Design.....	27
4. Literature Review	
a. Summary of findings - literature review.....	31
b. Documents reviewed.....	34
5. Survey	
a. Summary of findings - survey.....	43
b. Survey questions.....	50
6. Interviews	
a. Summary of findings - interviews.....	56
b. Interview guides (agency representatives/ program managers/grantees).....	63
7. Additional Resources and Background Readings.....	70

Executive Summary

Within the growing and evolving field of Global Health Research for Development (GHR4D), there is concern particularly among funding agencies, that robust, efficient and relevant evaluation strategies for this field are not readily available. A commissioned 14-week study was undertaken by the Canadian Coalition for Global Health Research (CCGHR) to identify and highlight current approaches, practices and challenges in the experience of funding agencies and funded researchers. Information was obtained through interviews and surveys. Supplemental insights were gleaned from document reviews of both peer-reviewed and non peer-reviewed literature.

The **key messages** from this study are:

1. **The context of global health research itself is evolving and changing.** From an earlier focus on health research needs and capacities in low-income countries, there is growing recognition of complex global challenges that affect all countries. These challenges require special approaches for multi-disciplinary and inter-sectoral knowledge production and application. As national health research systems particularly in middle-income countries become stronger, these countries are increasingly making a distinction between national and global health research agendas. Also, more attention is being paid to equity-oriented considerations and the role of innovation. Evaluation practices and approaches need to keep up with these changes, and if done well could strengthen the GHR4D field as it evolves.
2. There are helpful **evaluation practice examples** available. They come from several of the funding agencies contacted in this study. Other examples are country-focused or institution-based evaluation exercises. Some stories have been identified (appearing in boxes in the main document) that were selected to illustrate diversity and innovation of evaluation practices. Also included is an example of an inter-agency collaborative initiative focused on evaluation approaches and methods.
3. From the literature review several **innovative approaches to evaluating research and using evaluation findings** are described. Although most of these have been developed in high-income country settings, they could be adapted and tested with the evaluation of GHR4D in mind.
4. The **challenges** involved in evaluating GHR4D are considerable. Presented here are challenges that were commonly identified at all levels: funding agency representatives, program managers and researchers (grantees) and through all information sources: interviews, surveys and literature review. While details about these challenges can be found in the body of the report, the key points are summarized below.
 - The perception of complexity involved in evaluating GHR4D was wide spread. This pertained to the focus of evaluation--that is, health research for development that may have both expected and unforeseen consequences, and to the nature of GHR4D itself, for example, the multiple stakehold-

ers and several sectors involved. Global health problems, particularly involving health inequities are not only complex, but persistent, resistant to simple interventions and fraught with competing interests. The journey from research to changes in policy and practice, and particularly to health and societal impacts, is usually long, non-linear and multi-faceted.

- There was considerable variation in the understanding of the purposes of evaluating GHR4D. Often several purposes were mentioned, with the most frequent being accountability and advancing knowledge.
- A frequent observation was that wider and more appropriate stakeholder involvement was needed in designing and implementing evaluations, including stakeholders “on the ground”, researchers and other “users” of evaluation, including decision-makers. This observation reinforces the importance of context in designing evaluation strategies.
- There was some uncertainty about understanding and identifying evaluation approaches and frameworks. This was noted in particular among the grantees that were interviewed. An important insight is that evaluation approaches and frameworks need to be tailored to a particular context or situation, but with general principles that are transferable.
- Regarding methods and tools for evaluation, it was recognized that evaluations restricted to bibliometric measurements were inadequate, and that “mixed methods” were often needed. In general, the range of methods and tools cited was somewhat limited.
- The evaluation of impacts is perceived to be particularly difficult. However, several groups, identified in the literature review, have developed useful approaches (including taxonomies of impact categories) on this issue.
- Regarding the evaluation results which were actually used, we did learn of some examples from the interviews, although these were not numerous. Further examples were gleaned from the literature review. In addition, there were some analyses about the conditions where evaluation results were most likely to be used.

Additional study findings included the following:

- Some similarities and differences in evaluation practices and challenges were observed across the three levels of experience with Evaluating Global Health Research for Development (EGHR4D) - agency representatives, program managers and grantees. Similarities included the challenge in measuring scientific and societal impact. Differences included the finding that grantees expressed willingness to be involved in agency and program level evaluations, but were uncertain about whether their views were wanted. Grantees also expressed the desire for capacity building around research evaluation practices, but perceived limited investment in this area.
- There was insufficient evidence about the approaches used by different funding agencies and organizations to allow comparisons about convergences or divergences among them.

Derived from this analysis, the study team believes that there are no immediately available evaluation approaches that are “best suited” to evaluate GHR4D. We do however believe that there is **guidance** that could be usefully shared. This includes guidance about: identifying the purposes of evaluation, clarifying research outcome and impact categories, preparing an inventory of methods and tools along with guidelines for their appropriate use; suggesting steps needed to create and use context and user-specific

evaluation frameworks for specific situations; and drawing upon illustrative case studies, perhaps from a repository of examples.

In addition to these key messages, the study team puts forward the view that more evidence is needed about evaluating GHR4D. Sometimes known as “research on research”, this growing field of scholarship is in a relatively early stage of development, within which GHR4D could be included. Given the special features of GHR4D, there is a definitely room for more empirical work in this area. These special features include: an explicit equity-orientation, special attention to context (particularly in low resource settings), the complex nature of GHR4D problems, major capacity gaps between high and low income countries, and the need for longer-term strategies to evaluate social and economic development impacts.

* * *

There seems to be a remarkable opportunity, reinforced in many of the interviews, for more collaboration and sharing across agencies to address the challenges identified. Both GHRI and the ESSENCE group of agencies already represent collaborative platforms in this field, and indeed are taking steps to provide leadership. A clear message from this brief study is that this important collaborative initiative is much needed, and should be continued and intensified.

Introduction

Funding agencies that support global health research for development (GHR4D) are increasing in number and scope, supporting knowledge production in a diverse and evolving field of scholarship and practice. A particular challenge involves evaluating this investment in GHR4D—a challenge that includes making sense of results at program and organizational levels. It is critical to identify efficient, robust and meaningful evaluation strategies for GHR4D, not only to ensure accountability in funding practices, but also to use evaluation results for other purposes such as advocacy, analysis for learning and program change, and resource allocation.

This background paper is a contribution to an initiative of the Global Health Research Initiative (GHRI) in collaboration with the ESSENCE on Health Research group of agencies. The paper summarizes the findings of a 14-week study undertaken by a team of the Canadian Coalition for Global Health Research (CCGHR)¹. The study explored current practices, approaches and challenges involved in evaluating GHR4D. We did this by reviewing the relevant literature and interviewing representatives of funding agencies as well as funded researchers. The paper was discussed at a two-day invitational workshop in Ottawa, Canada (June 2-3, 2014) and then revised on the basis of feedback from GRHI and ESSENCE colleagues and other workshop participants. Both the paper and the workshop discussions will contribute towards the production of an “ESSENCE good practice document” planned for 2015.

In this report, we describe the current “lay of the land” about evaluating GHR4D, along with our analysis of the situation and some considerations for further dialogue and debate. We provide examples of current evaluation practices and the challenges identified, a synthesis and a discussion. There are also relevant attachments (appendices), some of which are summaries of the findings from the various information sources.

For the purposes of this project, the study team adopted working definitions of the key concepts of research, of global health research in particular, and of evaluation. These definitions are summarized in Attachment 1.

1 The CCGHR is a Canada-based not-for-profit global network committed to “*promoting better and more equitable health worldwide through the production and use of knowledge*”. [See: www.ccghr.ca]. The members of the project team included: Vic Neufeld and Donald Cole (co-leads); Alan Boyd and Donald Njelesani (research officers), Imelda Bates and Stephen Hanney (consultants), and Dave Heidebrecht and Roberta Lloyd (support staff).

The GHR4D context

Before we can discuss the evaluation of GHR4D, we need some understanding and hopefully consensus, about what GHR4D actually is—how it began and how this field is changing.

History: The GHR4D field has evolved over the past twenty-five years, dating back to the landmark 1990 report of the Commission on Health Research for Development. The stated aim of this report, entitled *Health Research: Essential Link to Equity in Development*, was: “...improving the health of people in developing countries”. The report also said: “Research should not be limited to the health sector, but should also examine the health impact of development in other sectors, and the socioeconomic determinants of health ...”. So, the GHR4D “movement” began with a focus on developing countries and included a broad view of health research for development.

This report had a remarkable impact over the next decade, including the creation of new organizations such as the Council on Health Research for Development (COHRED) in 1993, and the adoption of “essential national health research” (ENHR) strategies and structures by a significant number of developing countries. In 2000, the term ENHR was replaced by “national health research systems” (NHRS), still with a focus on developing countries (later to be more commonly known as “low and middle-income countries” or LMICs). The NHRS concept includes issues such as, governance, national health research policies and strategies, and national health research priority setting and implementation. Also about this time, the term “global health research” began to be used to capture the broad idea of health research for development.² An example is the creation of the Global Health Research Initiative (GHRI) in 2001.

Over this time, the global burden of illness has also changed. While communicable diseases continue as a major challenge, highlighted by the HIV/AIDS epidemic, along with persistent tuberculosis and malaria, the burden of non-communicable diseases has increased, particularly in low and middle-income countries (LMICs). The WHO Commission on Social Determinants of Health (CSDH) made three over-arching recommendations: improve peoples’ daily living conditions; tackle the inequitable distribution of power, money and resources; measure and understand the problem and assess the impact of action (2008, WHO CSDH report).

These problems and other challenges have been addressed in a series of major quadrennial events that brought the global health research community together-- in Bangkok (2000), Mexico (2004), Bamako (2008) and Cape Town (2012). To some extent, these events marked steps in the evolution of the field of global health research. For example, the 2004 “Ministerial Summit” in Mexico focused on research to strengthen health systems and on the role of knowledge translation. The Bamako event (2008) introduced the term “research for health” to emphasize the importance of trans-disciplinary research to improve the health of societies. Forum 2012 in Cape Town was entitled: “*BEYOND AID - research and innovation as key drivers for health, equity and development.*” A significant feature of this event was the strong voice and leadership of “the south”—particularly of African colleagues.

Current trends: Several features in this evolution of global health research (GHR) are worth highlighting:

- A growing appreciation of the interconnected **global** nature of GHR, recognizing that many important challenges that require the production and use of knowledge are indeed global—that is they affect all nations, “north” and “south”. Examples include the health impacts of climate change, human migration, emerging infections, trade agreements, antimicrobial resistance and others. Julio

2 Note: we use the term “global health research” (GHR) throughout this document since it is more commonly used than “GHR4D”, but with the understanding that GHR includes the “for development” concept.

Frenk and colleagues (Frenk et al, 2014)³, have recently stated that: “*we need to globalize the concept of global health*”. They argue for the inclusion of two key ideas: global health as the health of the global population (and not “foreign health”); and global health as a manifestation of health interdependence (not dependence).

- The fact of persistent **health disparities**. Although an equity-orientation has been a feature of GHR throughout this time, initiatives such as the CSDH have reinforced the central importance of equity-driven health research (2011, Ostlin et al).
- As part of an inter-sectoral trend, a focus on **innovation** has become a significant feature in this field. Some of this was launched by the early call for solutions to “grand challenges” issued by the Gates Foundation, and picked up by Canada’s Grand Challenges Canada program. A feature of this development has been more discussion about the role of the private sector (industry) in GHR. COHRED has embraced the innovation theme, and has announced its Forum 2015 with the title: “*People at the Center of Research and Innovation for Health*”. The announcement about this event claims that it will “*complete a transformation from redistributing resources from ‘North’ to ‘South’ (as in closing the ‘10/90 Gap’) to a new Forum that will emphasize the new global reality of greatly increased capacity, funding and potential of LMICs themselves, in terms of research and innovation.*”
- With the evolution of stronger national health research systems in middle-income countries (MIC)s, leaders in these countries are addressing questions such as: when does global health research become national health research (and vice-versa); what is the most effective interaction between **national** and **global** health research agendas from a country perspective. And when several MICs collaborate in their research endeavors, is this a new form of international health research?

What does this changing context mean for the evaluation of GHR4D? As a basic requirement, evaluation practices and approaches need to keep pace with these changes. If done well, evaluation practices can strengthen GHR4D as the field continues to evolve.

3 This reference and that of other articles in the body of this report are referenced at the end of the main document. Further references are provided in two attachments: 4b and 7.

Research Design

The need for the study was identified by the ESSENCE for health research group of funders with the aim of “taking stock of current practice and ways to improve the evaluation of global health research for development”. [More details about the research design can be found in Attachment 3: Research Design.]

Reflecting the guidelines in the original request for proposals (RFP), the study team identified four research questions:

1. *How are research funders and grantees evaluating global health research for development (GHR4D), with respect to approaches and frameworks?*
2. *What successes, best practices, challenges and solutions are being used to evaluate GHR4D at the organizational, program and project levels?*
3. *Are the identified approaches converging or diverging, and how?*
4. *What evaluation theories and frameworks are best suited to evaluate GHR4D?*

Information to address these questions was obtained from three sources: literature reviews, an electronic survey and interviews. Early in the project, the CCGHR study team and the funding advisory group (GHRI and ESSENCE members) discussed which agencies and organizations should be invited to participate in the study, leading to the identification of ten agencies. As the project proceeded, some changes in this list were made, as a result of some agency-specific considerations (see Attachment 2 for agency and interviewees in the four categories by each agency/organization).

Counts for each source were as follows:

- Literature reviews of peer reviewed and non-peer reviewed publications, as well as a review of relevant background documents—all related to EGHR4D. In summary, 156 documents and websites were judged to be relevant based on their title, among which: 35 were empirical studies/evaluations of GHR4D with some assessment of impact; 13 included descriptions or reviews of frameworks, approaches or methods for evaluating research; 58 were of general background interest and 26 were on health research capacity strengthening evaluation; and 24 were excluded as not relevant or text unobtainable (n=3).
- Electronic survey sent to researchers identified by participating agencies, as well as individuals associated with various global health research and evaluation networks and organizations (48 complete and 54 partial responses received—that is 102 responses out of more than 400 emails sent);
- Interviews with agency and organizational representatives (n=8), managers of programs within those agencies (n=2), regional health organization representatives (n=2), and researchers (grantees, n=4) recommended by agency or program staff and an evaluator (n=1). One agency representative and one grantee also had evaluation expertise—thus three evaluators completed interviews. A total of 17 individuals were interviewed.

As the data from the above sources came in and the analysis proceeded, on several occasions, the study team teleconferenced with members of the funding advisory group to review progress, and to obtain feedback on earlier drafts of the emerging background paper.

Findings

Summaries of the findings from each of the separate sources of information can be found in the attachments:

- Attachment 4: Literature review (summary of findings from 20 articles about GHR4D, and a table displaying features of the articles reviewed);
- Attachment 5: Survey (summary of findings from 48 complete responses and 54 partial responses), and the survey questions;
- Attachment 6: Interviews (summary of findings from 17 interviews, and the interview guides used in conversations with agency representatives, program managers and researchers).

Below we summarize overall findings; examples of current evaluation practices; findings from evaluation reports about approaches to evaluating GHR4D and using evaluation results; an analysis of some similarities and differences across levels and countries; and challenges experienced.

A. Current Practices:

The most important purposes of evaluation were generation of new knowledge, accountability, strengthened research capacity and informed policy almost equally among survey respondents. Interview respondents mentioned several purposes, among which advancing knowledge and accountability were the most important. Actual evaluation practices varied from relatively straightforward monitoring in relation to plans (a national research agency) through evaluations guided by theories of change (many program and project evaluations) to meta-level strategic evaluations (funder as a whole). Use of logical framework was common at all levels, while evaluations of projects by external consultants paid for by funders were less common. Summarized below (Table 1) are key interview findings at the levels of agency, program and grantee (project) about evaluation purpose, results sought and types of evaluation done. Some of the statements are what interviewees actually said.

Table 1. Evaluation purpose, results sought and methods

Level	Purpose - What trying to do with evaluation	Key Results Sought from Evaluation	Reported Types of Evaluation Undertaken
Agency	<ul style="list-style-type: none"> • Advance knowledge • Strengthen research capacity • Inform investments in health research • “know what actually research does” • Determine whether partnership was worthwhile? • Accountability to tax payers 	<ul style="list-style-type: none"> • Strengthened health systems across countries and regions • Evidence to support decision-making • Changes in individuals, institutions and national research systems • Knowledge translation – “collaboration with relevant stakeholders in order to optimize that the knowledge will actually be used by those stakeholders” 	<ul style="list-style-type: none"> • Monitoring and Evaluation processes • Concurrent monitoring, for the purposes of “learning while doing” • Logic models with performance assessment • 5-year external agency/program review • SWOT Analysis • Impact/ Cohort Analysis • Tracer studies of people trained • Portfolio analysis
Program	<ul style="list-style-type: none"> • “something answers the question, is this working, might this work?” • “learning for ourselves ...what approaches work well and what do not” • modifying programs and projects • advancing the field • Comply with organizational policies 	<ul style="list-style-type: none"> • The extent to which equity issues were uncovered and equity was promoted • Consequences of trying new approaches 	<ul style="list-style-type: none"> • Mid-term and final program evaluation • Results based management (RBM) • Rolling program reviews • Program outcomes review / Outcome mapping • Self-assessment reports • Impact analysis • Traditional evaluation (before & after) with emphasis on processes and outputs following a theory of change • Stories • Most significant change tool
Grantee/ Project	<ul style="list-style-type: none"> • Demonstrate effectiveness or impacts as part of research • Accountability (to donors) • “inform the next stages or project appraisal document or how next project is set up” 	<ul style="list-style-type: none"> • Equity impacts of a particular intervention • Change in practice as a result of a project • degree to which policy was influenced, • the impact of training activities • Capacity development within a project 	<ul style="list-style-type: none"> • Logic models and performance measurement frameworks • Mid-term and final project evaluation • RBM • Impact Analysis using economic approaches • Participatory Action Research & Evaluation • Outcome mapping • Donor initiated 6-month progress report

Presented below are examples of evaluation practice from different types of organizations: funding agencies, national health research systems and research institutions.

1. At the agency and organizational level:

Most of the agencies contacted provided examples and documents about evaluation practices that they found helpful. These included WOTRO's Monitoring and Evaluation plan (Box 1), the work of the GRHI evaluation group (Box 2), the PEPFAR (U.S.A.) evaluation plan, Sida's evaluation of its Zambia program, the IDRC Evaluation Unit's self-assessment and others. We provide three examples with publically available and/or peer review documentation from this group, to illustrate a range of experience:

Box 1

WOTRO case study: Taking Account of Research Funding Impact

WOTRO (the Netherlands) carried out an impact evaluation of WOTRO-funded research from 1998. The purpose of this cohort analysis was to explore the scientific and societal impacts of WOTRO-funded research. A specific pilot ex-post evaluation of the performance of one cohort of projects was conducted on projects completed in 2008, covering a time frame of five to seven years after completion of the projects, a time period deemed long enough to assess their impact. Of 54 projects funded by WOTRO in 1998, information on 40 projects was used in the analysis. The analysis focused on expected impacts in four broad categories: (1) scientific outputs, (2) impacts on future research and scientific capacity strengthening, (3) societal diffusion, dissemination and collaboration, and (4) impacts on policy and practice.

The analysis identified a number of factors that contributed to enhanced scientific and societal impacts. For example, it was observed that project outputs are realized mainly after the funding period, with key publications appearing, not within the first year after the project ended, but rather about 4-5 years after project funding has stopped.

In one instance, cohort analysis revealed that a study with a very narrow focus on health research actually had considerable development relevance even though during the funding period the project had little relevance and connection to broader development outcomes. The cohort analyzed was funded 15 years prior. Here's a quote from the analysis: *"We noticed that almost all the research we funded eventually turned out to be quite relevant for development, even though the research was not assessed for development relevance during the project period."*

WOTRO has used the findings of this evaluation to change its strategy. For example, engagement of relevant stakeholders outside the research community in all phases of the research project is now a requirement.

Source: http://nwo.nl/nwohome.nsf/pages.NWOP_899FTQ_Eng

Box 2:

Evaluation systems and processes that go beyond program and project evaluations: GHRI cross-program and “platform” evaluation

In recent years Canada’s Global Health Research Initiative (GHRI) has undertaken work to see how it can better assess value across research programs, and the added value that the GHRI “platform” provides to the programs and projects that it supports. This work has included an analysis of the impacts reported by projects from different GHRI programs, which highlighted inconsistent reporting of impacts, including qualitative indicators of impact and contextual barriers and facilitators of impact. The analysis suggested that these issues might be addressed through changes such as providing operational definitions for impact indicators and guidance on issues of attribution and contribution; and developing a sub-set of core indicators to be used for facilitating reporting across all programs.

An analysis has also been conducted of past evaluations and reviews of GHRI, finding that these do not amount to a platform-level evaluation, because they were not designed to be integrated and inter-locking. Gaps in the information available to assess GHRI as a platform have been identified, along with the need for further work to define impact and “rolled up indicators” across the three GHRI partner agencies. These analyses helped inform the development of GHRI’s latest evaluation plan.

MacLean, R. (2012) The Global Health Research Initiative: A preliminary review and analysis of potential cross-program impacts. Unpublished report. Canadian Institutes of Health Research. Institute of Population and Public Health.

Seifried, A. (2013) Taking stock of GHRI evaluations and reviews: what we have, what we need and suggestions for the future evaluation plan. First draft. Unpublished report. GHRI-IRSM.

2. At the national level:

Fewer country-level evaluation practices were identified. One of those that has been written up is the Guinea Bissau story (see Box 3).

Box 3

Evaluating a national health research system *

This paper represents a useful national evaluation case study. It describes how a health research system in a low-income country (Guinea Bissau) evolved and how it is currently performing. The team that conducted this evaluation, used the framework for a national health research system (NHRS) proposed by the WHO [See Pang T et al. Bulletin of the World Health Organization 2003;81:815-820]. They used a qualitative case-study method beginning with a document review and 39 in-depth interviews with a range of health researchers, policy makers and practitioners. Ten (10) research projects (3 ongoing and 7 finalized) were used as a diverse sample that would reflect the NHRS functions described in the framework. An iterative analysis approach was used, to determine key themes.

A key lesson was that the health research system and the health system itself co-evolve over time. Included in this observation is the substantial influence of the international research and development system on local research, resulting in relatively less attention to research questions relevant to local decision making. The paper also describes the co-evolution of research practices and systems that have taken place within the country, with a growing realization of the importance of local ownership of the NHRS, reflected in this sentence: “Ultimately, to achieve a sustainable NHRS, a continuous dynamic has to be realized within the country through which local priorities and funding leads to local research that leads to local action.” The authors observe that local efforts to develop a well-functioning NHRS may actually be constrained by international research and development influences, because of on-going dependence on external funding and other influences.

* Kok et al. Health Research Policy and Systems 2012, 10:5 [also Reference 2, Attachment 3]

3. At the research institution level

A published example from a research institution comes from Bangladesh. (See Box 4).

Box 4

Using a monitoring and evaluation framework: an institutional example

The International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) is a well-known research institute that was established in 1960. Recently, this institution faced the challenge of how to manage the number and variety of partnerships and funding arrangements. Two problems in particular were identified: insufficient core funding to build capacity and support the infrastructure, and inability to direct research funds toward the priorities in the institution's strategic plan. With support from their funders, specific strategies were implemented to address these problems. After three years, a review was conducted using an agreed upon monitoring and evaluation framework (MEF), including indicators.

A mixed methods approach was used during independent annual reviews carried out between 2006 and 2010. Quantitative data included the number of research activities related to stated priority areas, revenues collected and expenditures. Qualitative data included interviews of ICDDR,B research and management, research users and key donors. The review revealed that changes made to funding arrangements, supported by an effective MEF, helped the organization to better align funding with research priorities and to invest in capacity building.

For details, see: Mahmood et al. *Health Research Policy and Systems* 2011, 9:31

<http://www.health-policy-systems.com/content/9/1/31>

B. Approaches and methods for evaluating research and increasing the use of evaluation results:

Several frameworks and methods for evaluating research have been devised, but most of those in the published and non peer-reviewed literature have been developed in high-income countries and have not been tested with the evaluation of GHR4D in mind.

In late 2012, the UK Collaborative on Development Sciences (UKCDS – a collaborative arrangement that brings together 14 UK government departments and research funders that work in international development), DFID and IDRC convened a workshop of experts to consider the challenges of evaluating the impact of research programs for development. They subsequently developed a resource based on the workshop (Thornton and Shaw 2013)⁴. Rather than providing a comprehensive assessment of methods, this resource highlights some potentially useful methods and some of their likely strengths and weaknesses. Components of this resource (available as .pdf documents from: www.ukcnds.org.uk) include:

- Why and what? Motivations for evaluating impact and tools for framing the right question
- Important issues to consider in evaluating a research program (best practice)

4 Note: the documents mentioned in this section (Approaches to Evaluation of GHR4D) can be found toward the end of Attachment 7: "Additional Resources and Background Readings".

- Approaches and methods for evaluating the impact of research
- Potential challenges to more effective evaluation of research impact
- A list of links to further resources

The HERG Payback, SIAMPI, NIHR 'lean'/dashboard approach and RAPID Outcome Assessment (ROA)/RAPID Outcome Mapping (ROMA) are cited as being frameworks for specifically evaluating research impact. Each is described briefly in the resource, with links to further information. The presentation describing the Payback Framework ⁵also explored how the framework could potentially be used to illustrate the rate of return on health research in low-and-middle-income countries, by combining it with an adaptation of the approach used in 2008 to show the high rate of return to UK health research ⁶. Various tools that might be used are also described, including some that have been used in a development context, such as *Contribution analysis, Outcome Mapping, Process tracing and Most Significant Change*. Alternative metrics to traditional bibliometrics, “Altmetrics”, are also highlighted. Tools cited for use in framing evaluation questions include Theory of Change, the Logical Framework Approach and stakeholder mapping. The authors state however that they are not aware of a framework for choosing between methods.

In the rest of this section we seek to augment this resource, by highlighting additional material that either updates information or addresses gaps in this document, such as support for choosing between methods, and for increasing the use of evaluation results.

New approaches, and ways of choosing between them

A method called *Impact Oriented Monitoring* (IOM) (Guinea, *et al* 2013), based on the Payback model, is being developed through the EVAL-HEALTH project (European Commission 2011) to evaluate the impact of funded research projects in public health in developing countries. IOM is developing various tools, including a project results framework, a coordinator’s survey and an end users’ survey and an assessment tool. Also of interest are reports that have compared different frameworks for evaluating health research, albeit without a focus on LMICs or on global health. (Brutscher, *et al* 2008) compares eight research evaluation frameworks in use in HICs. The frameworks are compared against five dimensions: evaluation objectives; outcome measures; levels of aggregation; timing and evaluation methods. The authors’ analysis suggests that the choice of evaluation objectives should be the basis for the choice of framework.

Further research by RAND (Guthrie, *et al* 2013; also see Box 5 -below) takes this analysis a step further through an investigation of a further 14 frameworks, concurring that to be effective, the design of a framework should depend on the purpose of the evaluation, which may be one or more of the following:

- **Advocacy:** to demonstrate the benefits of supporting research, enhance understanding of research and its processes among policymakers and the public, and make the case for policy and practice change

5 Hanney S (2012) The Payback Framework: developments in assessing policy and economic impacts of health research & application to development research. Available in: <http://ukcds.org.uk/resources/evaluating-the-impact-of-research-programmes>

6 Health Economics Research Group. Office of Health Economics, RAND Europe (2008) Medical Research: What’s it worth? Estimating the economic benefits from medical research in the UK. London: UK Evaluation Forum: www.wellcome.ac.uk/economicbenefits

- **Accountability:** to show that money and other resources have been used efficiently and effectively, and to hold researchers to account
- **Analysis:** to understand how and why research is effective and how it can be better supported, feeding into research strategy and decision-making by providing a stronger evidence base
- **Allocation:** to determine where best to allocate funds in the future, making the best use possible of a limited funding pot.

A practical guide to developing a research evaluation framework is provided in the form of a decision tree and 13 key questions for an organization to consider, together with an example to illustrate the approach. They also note that new evaluation approaches may be needed for emerging areas of research, such as implementation science.

Box 5

Guidelines for building research evaluation frameworks *

This monograph, produced by RAND Europe, was commissioned by the American Association of Medical Colleges (AAMC), with the charge: *to conduct a detailed synthesis of existing and previously proposed frameworks and indicators used to evaluate research* about “..how research outcomes can be measured in different context and ultimately account .. for returns on investment”. Included is a detailed analysis of six (6) frameworks, and a further analysis of seven (7) more frameworks. Of the 13 frameworks eleven (11) were from high-income countries and two from middle-income countries (Argentina and South Africa).

Based on this analysis, the final chapter in the monograph provides guidelines about how to build a new research evaluation framework. Particular emphasis is placed on determining the purpose of conducting the evaluation, suggesting four (4) purpose categories: advocacy, show accountability, analyze (to understand why research is effective and how it can be supported), and allocate. The executive summary highlights several key findings, the first of which is: *Designing a research evaluation framework requires trade-offs; there is no silver bullet*”.

* Guthrie S et al: *Measuring Research: A guide to research evaluation and tools*. (2013) RAND Europe

Approaches for increasing the use of evaluation results

A technically perfect evaluation is of little value if its results are not used or the evaluation process does not produce useful learning for those involved in it. The findings from our interviews (see section C) indicate that evaluation results do not always influence the strategic decision making of funder agencies. The underutilization of evaluation results by organizations generally has been noted as an issue for many years, with suggested solutions including better evaluations, getting closer to decision making, building up an evaluation culture and communicating better (Mayne 2014). Challenges to developing and building an evaluation culture within the World Health Organization (WHO) have been described (Santamaria, *et al* 2014) (cited by Mayne, 2014).

(Ramalingam 2011) lists insights from previous research that might help maximize take-up of impact evaluation results in an international development context:

- Institutional readiness: understand the key stakeholders; adapt the incentives; invest in capacities and skills
- Implementation: define impact in relation to the specific context; develop the right blend of methodologies; involve those who matter in the decisions that matter
- Communication and engagement: communicate effectively; be persistent and flexible

A multi-sectoral evaluation of HIV programs in Papua New Guinea suggests that close interaction between the commissioner and the evaluation team throughout the evaluation process is necessary to produce an influential evaluation (Rudland 2011). This not easy, and relies on resilient working relationships being built, a strong evaluation team leader and a sophisticated understanding of ‘independence’ in evaluation from both parties. Commissioners need to be more active than simply managing the contract, helping create a receptive organizational environment.

Taut and colleagues (2007) conducted action research to improve the ability of members of a large international development organization to learn from evaluation through self-evaluation, i.e. small-scale evaluation projects carried out by staff as part of their everyday work activities in order to answer questions concerning their work. The author concludes that supportive organizational culture and structures are needed if such initiatives are to have widespread, sustained effects. Findings that may be of interest to others planning similar initiatives are described.

Tennant (2010) explored how evaluation use and influence theories might assist the Australian Agency for International Development (AusAID) to realize the potential of evaluation, suggesting that there could be significant benefits from valuing and treating evaluations as interventions in themselves, in line with evaluation influence theory. Over time, learning could be accumulated about what works in enabling evaluations to be influential.

An analytical framework for improving the understanding and use of evaluation of humanitarian action (Hallam and Bonino 2013) may also provide insights for EGHR4D practitioners and policy makers as it draws on the wider literature on evaluation capacity development and evaluation use and makes some practical suggestions for improvement. The framework is hierarchical, with the most important and fundamental issues of leadership, culture, structure and resources appearing in Capacity Area 1. Clarifying purpose, demand and strategy are also important but less significant and so appear in Capacity Area 2. Capacity Area 3 focuses on processes and systems that, while useful in their own right, are considered less likely to bring about fundamental change on their own, without changes made elsewhere. A follow-up study is designing a self-assessment tool to help agencies reflect on their evaluation processes, take stock of their practice in evaluation utilization and uptake, and identify areas on which to focus future efforts.

Further resources

An implementation research toolkit (TDR WHO 2014) contains a module on monitoring and evaluation to examine the difference between the implementation effectiveness and efficacy of health interventions. Key steps in developing a monitoring and evaluation plan are described, so as to assess implementation outcomes: acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration and sustainability.

Kok and Schuit (2012) provide a detailed description of the four stages and ten steps of *Contribution Mapping*, an approach to evaluating research that is designed to be useful for both accountability purposes and for assisting in better employing research to contribute to better action for health. An indication is provided of how the approach can be used for the purposes of accountability, of learning, and of improvement.

Drawing on recent experiences of using ‘theories of change’ amongst organizations involved in the research–policy interface in an international development context, (Barnett and Gregorowski 2013) suggests that theory-based approaches provide a way forwards to understanding and measuring policy change and impact, given the difficulty of pre-planning change. Rather than trying at the end of a project to evaluate a depiction of change envisaged at the start, incentives need to be in place to regularly collect evidence around the theory, test it periodically, and then reflect and reconsider its relevance and assumptions. Questions to consider on an ongoing basis about outcomes/impact, about activities, and about attribution are suggested. Four key lessons are also identified.

With regard to evaluation of clinical research management, (Dilts 2013) present a “*Three-plus-one*” model that incorporates three local levels (individual research study; managerial; strategic) and a global/multi-institutional level. The area of focus at each level and potential metrics are suggested. This model may provide insights for the evaluation of GHR4D initiatives such as EDCTP and IAVI.

C. Similarities and Differences:

As we analyzed the findings across the three sources of information (literature review, surveys, and interviews), some findings spoke to the issue of similarities and differences. Across organizations, we noted differences in survey responses by the roles that organizations played in GHR4D. In comparison to organizations which were primarily funding or commissioning GHR4D, those organizations doing or evaluating GHR4D tended to place a higher priority on generating knowledge, building research capacity in LMICs, informing decision-making by health policy makers and producing health benefits in LMICs.

1. Are there differences in evaluation practices and experiences across the levels of agency, program and project (researcher or grantee):

The main difference among interview respondents was a greater focus on the equity agenda among program-level staff, project-level staff and evaluators, compared with agency-level staff. Inequities of interest included North-South partnerships, coverage of services and gender. Project level participants tended to use more participatory approaches. More specific findings are summarized in Table 2 (below) and the challenges elaborated in the next section:

Table 2. Challenges and Successes in evaluation

Level	Challenges	Successes
Agency	<ul style="list-style-type: none"> Measuring impact (scientific and societal) Prioritization of biblio-metrics over other types of evidence Understanding of how to use research to influence policy and decision making Use of results and reported activities to influence agency strategic decision-making. Inter-agency collaboration and cross cutting learning Limited involvement of grantees in agency evaluation /review 	<ul style="list-style-type: none"> Contribution of M & E to making India Polio Free Cohort analysis tracking grant recipients 10 years after end of funding Annual review crucial and push PI to go after goals. Key way to keep track of activities and investments Impact evaluation demonstrated impacts on health system that senior management needed for endorsement of continued funding
Program	<ul style="list-style-type: none"> Measuring impact (scientific and societal) Ensuring relevance of research questions to development/ country context Connection between reported activities (Program Managers) and strategic planning (Board) – Ensuring continued funding of activities. <i>Yearly struggle</i> Limited investment in evaluation capacity 	<ul style="list-style-type: none"> Cross learning through inter-agency evaluation working group (GHRI) Operational research and ability to apply lessons in different contexts Inclusion of grantees in program reviews
Grant/ Project	<ul style="list-style-type: none"> Measuring impact (scientific and societal) Donor and stakeholder relationship/engagement. Is grantee able to raise challenging questions? How receptive is donor or govt. rep? Ensuring relevance of research questions to development/ country context Limited investment in evaluation capacity 	<ul style="list-style-type: none"> Flexibility and willingness to consider alternative evidence beyond bibliometrics Participatory approaches involving key stakeholders Equity agenda

Although all participants recognized that bibliometrics do not constitute sufficient evidence for evaluation, they were perceived as useful for communicating the results/impact of interventions. Among different kinds of organizations, bibliometrics were regarded as particularly useful to NGOs for demonstrating credibility to funders and government stakeholders and enabling buy-in and a willingness to collaborate.

Program leads and evaluators recognized that community level research and NGOs found evaluation helpful to improve their interventions and generate different types of evidence needed. Significant change, testimonials and collective knowledge/shared experiences were cited as relevant evaluation methods. At the program level, participants also identified economic impacts as particularly useful for demonstrating the impact of their evaluations but noted the lack of capacity in this area. Participants perceived econometric analyses as key to influencing government decision-making.

2. Are the approaches converging or diverging?

There was insufficient evidence emerging about the approaches used by different funding agencies and organizations to allow statements about trends towards divergence and convergence. There was convergence both about the “thinking” about EGHR4D -for example, that it is a complex challenge and that the evaluation of impacts is difficult. With more groups becoming interested and involved in working in this area, more views are coming forward which may, at least temporarily, contribute to more divergence.

D. Challenges:

The challenges in evaluating global health research are considerable. They include the following:

1. A general finding, particularly from the interviews, was the perception of complexity in evaluating GHR4D. This pertained to the subject of the evaluation e.g. health services reforms and their unforeseen as well as foreseen consequences, the nature of global health research e.g. multiple stakeholders and often sectors involved, and the evaluation of global health research e.g. multiple implementation activities and outcomes. For some, global health problems such as health inequities are not only complex, but also persistent, resistant to simple interventions, and fraught with competing interests (Morrison et al, 2013).

Complexity poses challenges to devising a tractable, parsimonious theory of change for projects or programs. It makes attribution of the research outcomes to particular activities more difficult. The “Outcome Mapping” framework developed by IDRC’s Evaluation Unit can be helpful here, with its focus on outcomes carefully described by both the researchers and the groups with whom the research is being conducted. (Earl & Carden 2002).

2. There was considerable variation in the perceived purposes of evaluating GHR4D (see table 1 above). Sometimes confusion occurred between the purposes of evaluation initiatives per se, as distinct from the assessment of the achievement of the objectives of research projects. For example, one agency representative spoke of “*a shift from proving to improving global health interventions and programs*” but not all others voiced such a learning purpose. Some spoke of eagerly looking for unintended or unexpected outcomes—including honest accounts of “failure”, and the lessons learned as a result. Others were more fearful of such stories, as they might jeopardize funding to their programs. We heard stories from examples of inter-agency initiatives about tensions (regarding the purpose of evaluation) among participating agencies. One program officer said:

“We spent some time negotiating the purpose of our evaluation plan, with one partner concerned primarily with accountability (“return on investment”) and another on learning for the purpose of program improvement.”

Several of the resources in the non peer-reviewed provided guidance about clearly defining the purposes of evaluation (see the paragraph above about the RAND study—Guthrie et al, 2013).

3. Interview, survey and workshop participants observed that wider and more appropriate stakeholder involvement was needed in evaluation for GHR4D. This included the recommendation for more participatory involvement of people “on the ground”, as well as more involvement of decision-makers. Essentially, this is another reflection of the importance of context. Particularly in the field of global health research development where equity is a key underlying value, it is important to obtain the views of potentially marginalized stakeholder groups. The literature is replete with information on the importance of the involvement of stakeholders at all levels and throughout the evaluation and project life-cycle, and this was reflected in long serving agency and program representatives comments. The voices and perspectives of donor recipients were seen by most agency and program representatives as relevant to informing the short and long-term donor strategies and priorities of donors. The omission of these voices and perspectives can further the perception of evaluations as exercises in fiscal accountability primarily for donors (see box 6).

Box 6

Stakeholder involvement for whom?

A representative for an organization (LMIC) funded by an international donor shared the experience of how their relationship evolved over time with successive health projects increasing in scope (financial, type) over approximately 10-year period. Some notable changes have included the development of evaluation capacity particularly around the understanding and application of *theory of change* within the organization's evaluation approaches with relatively a more developed culture of evidence-based decision-making being observed.

As funding has increased over the period so too has the frequency of evaluations (end of year; mid-term) by the donor. Asked if and how their organization had participated in the evaluation of the donor, the representative noted *we have not participated in or seen any evaluation report of the donor, it would be good to one day see an evaluation of them*. The participant understood that not all donor recipients can participate in donor evaluations but they did believe recipients should at the minimum have an opportunity to inform donor evaluations and access to completed evaluations especially after a 10-year relationship.

4. There was some confusion among interviewees about evaluation frameworks. Some interviewees when asked to describe what framework was being used needed clarification about what this term meant. In fact, even in the literature review, there is some confusion. For example, what's the difference between a "theory", an "approach", a "framework" and a "model"? In fact, for this study, the team settled on the term "approach" since this seemed to represent an overall "umbrella" term for this concept. We also used the term "framework" if there was some specificity, perhaps displayed in a diagram in a document.

One program representative noted:

"To be able to tell or show funders how money was used, an important element is learning for ourselves [at a strategic level] what approaches work well and what does not. But also for our project, we want to keep an open mind and reflect on their approaches."

An important insight is that approaches and frameworks need to be tailored to a specific context or situation. Some useful guidance about how to build research evaluation frameworks is available (See Box 5 above).

5. There were a number of issues related to methods and tools for evaluation. For projects, one agency representative noted *"In the past you had designs that required fidelity, now you need experimentation and flexibility"*. Increasing rigour as valid ways of evaluating interventions, are a foundation for evaluation of research programs. As one interviewee noted:

"Credible evaluation includes some quality of research, and also connection to users of research, for example, what was helpful for enlightenment in knowledge use and grantee perspectives and outcomes."

In relation to quality of the research, several funders still include, but tend to move away from bibliometric measurements as a primary indicator of quality. A survey respondent said:

“The orientation around evaluation must move past impact, towards more qualified, nuanced understandings of research evaluation, such as realist evaluation and other theory-driven approaches, which are necessarily date, time and resource heavy, which often runs counter to granting time cycles.”

An agency representative commented:

“We have moved away from wanting to count the number of publications. We want to focus on what has changed and how our support led to some change for the better.”

Tensions were noted between simpler results-based management approaches, in which the data collection burden is limited but falls primarily on the grantee, and a broader range of mixed methods of evaluation. The latter may require additional resources, and training of researchers and implementing institutions. Several agency representatives spoke to this issue. Interestingly, methods cited in the peer-reviewed literature were somewhat limited, similar to evaluations of health research capacity strengthening (Cole et al 2014). This finding leads to the suggestion that an inventory of available methods and tools might be useful (building on the UKCDS materials cited above), along with guidelines about their appropriate application.

6. The evaluation of impacts was perceived to be particularly difficult. This was a finding from all information sources and across all levels (agency, program and project). Frequently reported was the observation that to evaluate impacts properly, more resources and longer time frames were needed. A survey respondent commented:

“A more consultative, flexible approach by the funder, which is not driven by the need to ‘show results’. At times there are no results, i.e. the evaluation is ‘empty’, and this needs to be acceptable to funders as well. Timelines need to be longer, and evaluative accountabilities need to be multi-polar, not simply oriented to the funder in order to capture all the impacts and processes of whatever the intervention has introduced.”

A practical approach to this dilemma might be to think about two kinds of impact: relatively short-term (within the time frame of a given project) and longer-term (requiring additional resources and effort). We did find useful guidance about evaluating impacts in the literature review. An example is the HERG “Payback framework” (mentioned in the Discussion section below).

7. Obviously, a well-done evaluation is of little value if its results are not used or if useful learning by those involved is minimal. We had expected to hear more about using evaluation results. We did find some examples where evaluation findings were used to improve processes and outcomes, but these were not numerous. One example is the WOTRO experience (Box 1 above) where the findings of a longer-term follow-up study were used to change the agency’s evaluation strategy. Another example, brought to our attention by an agency representative, describes how a systematic evaluation process contributed to a specific health outcome at a national level (see Box 7 below).

Box 7

Using evaluation results – an example

Here is an example, mentioned by an interviewee, of how a well-designed and conducted monitoring and evaluation initiative contributed to India's certification as a polio-free region.

In March 2014, the World Health Organization (WHO) certified India and 10 other countries in the South East Asia Region as Polio-free regions. The Global Polio Eradication Initiative points out that with a population of over one billion people, India was once considered the most challenging place on earth to end polio. Attaining the polio-free status was a result of a number of factors including commitment at all levels of the health sector, technological innovations, 2.3 million vaccinators, domestic financial resources and close monitoring of the polio program which led to immunization levels to rise to 99 per cent coverage.

The polio monitoring and surveillance system has been identified as a vital component in India's success. The polio surveillance system is now helping build capacity for India's Universal Immunization Program. *“Real time tracking e.g. quarterly tracking of a set of indicators provides data that helps governments understand whether they are on the right track and where they can adjust programming... how can we know in the next 6 months what is showing promise? Ongoing real time monitoring, learning and evaluation (MLE) is helpful here... monitoring and evaluation played an important role in India becoming polio-free.”*

In addition (as described in Section B above), the literature review found several examples of how evaluation results were used. Included in the research about this issue, are some determinants to maximize the uptake of evaluation results. These include a determination of institutional readiness, the appropriate involvement of stakeholders in the evaluation process, and the need to be “persistent and flexible”

E. Other findings:

An important finding from the survey and from interviews was the view that more collaboration and knowledge sharing about evaluation was needed among funding agencies. In fact, there are active examples of inter-agency collaboration. The commissioners of this study represent a strong example, since the commissioner group represents a joint initiative of two collaborations: the Global Health Research Initiative (GHRI) that involves three Canadian federal agencies, and the ESSENCE on Health Research group that currently includes 26 member organizations from around the world, all of whom are committed to promoting and supporting global health research for development. It would seem that most survey respondents (and some interviewees) are not aware of these important examples of inter-agency collaboration, so perhaps some pro-active “publicity” is needed.

An example of inter-agency collaboration relating to evaluation of GHR4D can be found in box 8:

Box 8

Inter-agency Collaboration on Evaluation

INDEPTH is a “southern-led” network of research centres that conducts health and demographic surveys of geographically-defined populations in low and middle-income countries (LMICs). Several funding agencies, including Sida support this network. As part of its policy to regularly evaluate organizations that it supports, Sida began to plan an evaluation of its investment in 2009. To avoid duplication of effort, Sida consulted with other funders about the terms of reference for this evaluation, so that the results could be used by all the supporting funding agencies. Also consulted was the INDEPTH secretariat who indicated that the organization would benefit from an external review to help assess the network’s performance and to contribute to the creation of its next strategic plan. The evaluation thus covered all of INDEPTH’s activities, extending beyond those supported by Sida. In addition to sharing the report with the organization, it was also shared with all funding agencies involved.

Reference: SIDA Review 2010:11. Reviewers: Samson Kinyanjui and Ian M. Timaeus. ISBN: 978-91-586-4127-3

Discussion

The study team recognizes several limitations to our study. Among them are the following:

- We were unable to arrange interviews with agency representatives, particularly some agencies in low and middle-income countries – for example, the NRF in South Africa and the NIMR in Tanzania. It proved to be difficult and time-consuming to find the most appropriate person to interview.
- Whenever possible, we requested that agency representatives provide examples (and documentation) of evaluation approaches and frameworks used. We also requested names of those managing program portfolios and, within those, grantees that we could contact for the survey or interview. Provision of such contacts by agencies was less common than we had anticipated. Hence our overall numbers of interviewees was smaller than anticipated. Survey response rates were likely low due to any or all of: the limited attention to the topic, the relatively specialist nature of the field, the short time line provided, or the lack of incentives for responding.
- Early in the project there was a discussion between the commissioners and the project team about which agencies to include in the study. The study team urged the inclusion of agencies from low and middle-income countries—a suggestion that was accepted. This led to a balance of agencies based in high-income countries, and those based in low and middle-income countries. However, in the design of the project, we did not include questions that might provide evidence about whether there might be differences in evaluation approaches and methods used between these groups. For example, no question was asked on the survey nor posed to interview respondents on comparisons between HIC vs LMIC agency approaches.

Nevertheless we think our findings are sufficient to prompt discussion on the following general issues about evaluating GHR4D:

Researching more Options for the Evaluation of GHR4D Initiatives, Programs and Projects, Processes and Impacts

Although we found reports in the literature and heard stories from interviewees about promising examples of evaluation of GHR4D, we as a project team realized how limited was the evidence base to support evaluation guidelines. More rigorous work is needed examining and reporting on the processes by which impactful GHR4D is designed, implemented and shared with relevant audiences. Such work should ideally consider the different national and global contexts in which GHR4D occurs, at the different levels project, program, institution-wide portfolios. The former may be more important for researchers, and the latter more for agencies, national and international, but we were struck by the overlap in interests in such evaluation work across levels. Building this field could coordinate with the broader field of evaluation of development initiatives (or “research on research”), to include evaluation of capacity building, and impacts of health research funding. We were glad to come across the EU-funded EVAL-HEALTH project (see: www.eval-health.eu) aimed at strengthening M & E practices of EU-funded interventions in developing countries. One of the project objectives is to “*develop and test a new methodology to monitor and assess the impact of research projects ...*”. We spoke directly with this group, and they are keen to collaborate in some mutually beneficial way building the evidence base for evaluation of GHR4D.

Considering how much evidence is needed to proceed with providing guidance

This study addressed the question: “*What evaluation theories and frameworks are best suited to evaluate EGHR4D?*” The short answer probably is that there are no immediately available evaluation frameworks that are “best suited” for GHR4D. Reviewers were mixed in their judgment on the evidence available, indicating divergence on the criteria for “good” or “better” practices in relation to EGHR4D.

However, the study team believes that there is sufficient available evidence and experience to strengthen current evaluation practices.

We offer some further considerations on this issue:

1. Distinctive features of GHR4D—implications for evaluation:

The overarching challenge of the global health research “system” is that it is inequitable, in terms of investments and capacities. Most of health research investment globally is still targeted on improving the health of societies in high-income countries. In addition, as noted in the “Context” section at the beginning of this document, the field itself is changing and evolving. Thus, the evaluation of GHR4D needs to consider realities such as the following:

- There are major resource differences that separate (most) GHR funding sources and their recipients, particularly researchers in low-income countries. As a consequence, evaluation strategies need to pay special attention to the context in which both the research itself and the evaluation of that research is conducted. Considerations should include the changing socio-political climate, and the research priorities identified by partner institutions and countries.
- The challenges being addressed in GHR are particularly complex. These complexities disproportionately affect vulnerable populations in low-income settings--examples include the health impacts of climate change and human migration. It is therefore critically important that these complex problems are addressed by trans-disciplinary teams, and involve the participation of relevant stake-

holders. Ideally, these stakeholders should include policy makers from several sectors, to ensure multi-sectoral policy influence, likely through a group of related policies.

- There are major research capacity differences at various levels: individual, institutional and national systems, between high and low-income countries. Evaluation approaches therefore need to recognize the capacities of national health systems (including health research systems) to both produce and use evidence. In addition, the capacities to conduct relevant evaluations and use the findings will need to be strengthened.
- The fact of persistent inequities (and the underlying “value” that explicitly addresses equity) needs to be recognized in ethical standards that guide GHR4D and its evaluation.
- We are concerned with global health research for development, which implies attention to longer-term societal benefits—both social and economic, as well as longer-term institutional and “system” strengthening. The evaluation of GHR4D should therefore reflect this “development” reality, for example by conducting several linked evaluations at different points in time.

Because of realities such as these, effective, respectful and sustainable research partnerships (including both “north-south” and “south-south” partnerships) are particularly challenging to develop and nurture. Similarly challenging are the arrangements for accountability—to whom are research partnerships primarily accountable—the funders, institutions in the country where the research is being conducted, the people in these countries? These same considerations apply to evaluation practices—who should be involved in designing and conducting an evaluation initiative, and who “owns” the evaluation results?

These realities and tensions need to be better understood and tackled, for example examining “insider-outsider” perspectives, short and long-term impact evaluation, and addressing both learning and accountability evaluation processes.

2. Practical guidelines are available:

From both the experiences offered by colleagues that were interviewed, and also insights from the literature review, we propose that there are general guidelines about the evaluation of GHR4D that could usefully be captured and shared. These include:

- Clarifying the purpose of a given evaluation initiative—for example, drawing upon the four “A’s” proposed by the Guthrie study: accountability, advocacy, allocation and analysis (for learning and improvement).
- Proposing a frequently used set of research impacts that can be used and adapted. An example is the Health Economics Research Group (HERG) Payback framework that categorizes the types of “payback” (that is, benefits from research) into the following five domains: advancing knowledge (that is, knowledge production); research targeting and capacity building; informing policy and product development; health and health sector benefits including improved health and increased health equity; and improving broad social and economic (development) benefits⁷. Originally used to examine the impact of health services research in the UK, the framework has also been applied to other areas of research (e.g. social science research), and includes a model that facilitates application to a wide range of types of research, in particular research in which potential users collaborate with

7 The payback framework is described in depth in: Hanney SR, Grant J, Wooding S, and Buxton MJ. Proposed methods for reviewing outcomes of health research: The impact of funding by the UK’s Arthritis Research Campaign, *Health Research Policy and Systems*. 2004 2:4 <http://www.health-policy-systems.com/content/pdf/1478-4505-2-4.pdf>

researcher on setting agendas. A range of organizations have used the model, including the Canadian Academy of Health Sciences study on “Return on Investment”, and ICDDR,B (Bangladesh).

- Preparing an inventory of methods and tools, along with guidelines for their appropriate use;
- Recognizing that any evaluation approach is context specific and “user” specific. There is no “silver bullet”, rather there will always be trade-offs. For most organizations and groups, this means that situation-specific frameworks and plans will need to be prepared. Guidance about how to create and use these situation-specific evaluation frameworks would seem to be useful and relevant.
- Drawing upon the many helpful examples (case studies) that are available, including lessons learned--what worked or didn't work.
- Guiding evaluation of GHR4D at different levels through a series of questions, which a multi-stakeholder group could work through to design and evaluation.

* * *

There seems to be a remarkable opportunity, reinforced in many of the interviews, for more collaboration and sharing across agencies to address the challenges identified. Both GHRI and the ESSENCE group of agencies already represent collaborative platforms in this field, and indeed are taking steps to provide leadership. A clear message from this brief study is that this important collaborative initiative is much needed, and should be continued and intensified.

Selected References

Note: Some readers of this document may wish to have the references to journal articles mentioned in the main document. For ease of access to these articles, these references are listed here. As indicated in the document, more extensive listings of references can be found in the Attachments:

- Attachment 4b: Documents (both peer-reviewed and non peer-reviewed) that were selected and analyzed with a focus on impact assessment of GHR4D.
- Attachment 7: Lists additional resources and background readings.

Cole, D., Boyd, A., Aslanyan, G. & Bates, I. (2014) Indicators for tracking program to strengthen health research capacity in lower- and middle-income countries: a qualitative synthesis. *Health Research Policy and Systems*, **12**, 17

Frenk J, Gómez-Dantés O, & Moon S. (2014). From sovereignty to solidarity: a renewed concept of global health for an era of complex interdependence. *Lancet* **383**: 94-97

Kok MA., Rodrigues A., Silva AP., & de Haan S. (2012) The emergence and current performance of a health research system: lessons from Guinea Bissau. *Health Research Policy and Systems*, **10**:5

Mahmood S., Hort K., Ahmed S., Salam M., & Cravioto A. (2011) Strategies for capacity building for health research in Bangladesh: Role of core funding and a common monitoring and evaluation framework. *Health Research Policy and Systems* **9**:31

Morrison V., Keeling M., Jacques M. Wicked Problems and Public Policy. National Collaborating Centre for Healthy Public Policy – Institut national de santé publique du Québec. Montréal (Québec), 2013. Accessed 2014 June 17 at http://www.ncchpp.ca/130/publications.ccnpps?id_article=927

Ostlin P, Schrecker T, Sadana R. and others: (2011). Priorities of Research on Equity and Health: Towards an Equity-Focused Health Research Agenda. *PLoS Medicine*, **8** (11): e1001115. Doi:10.1371/journal.pmed.1001115

Attachments

Attachment 1 – Definitions

Some definitions

For the purposes of this project we adopted a broad definition of health research: “the [production] of knowledge with the aim of understanding health challenges and mounting an improved response to them. This definition covers a spectrum of research, which spans five generic areas of activity: measuring the problem; understanding its cause(s); elaborating solutions; translating the solutions or evidence into policy, practice and products; and evaluating the effectiveness of solutions.” (Terry and van der Rijt 2010).

We also used a definition of global health research, derived from a definition of global health (Koplan et al 2009), and an analysis of global health research case studies (Boutilier et al 2011). With an overall commitment to health equity, the features of global health research include: long-term sustainable North-South partnerships, inter-disciplinary responses to complex issues, grounding in local contexts, and an orientation to policy or practice impacts.

We also adopted a fairly broad definition of evaluation. This includes process and impact evaluations, evaluations that aim to produce learning for or about a project, and evaluations that aim to judge worth or value. Like research, evaluation is deliberately planned and designed activity that ideally gives consideration to issues of ethics and of generating valid knowledge, according to appropriate standards of validity for the type of evaluation being conducted.

Evaluation of research may be conducted by external or internal agents. It can be conducted for a variety of purposes, but typically with a view to informing decision-making or planning – strategic and operational – and ultimately improving the research that is conducted (increasing impact, greater efficiency, etc.). A health research funder or grantee organization might conduct evaluation at a number of levels, including: evaluation of individual projects; evaluation of whole research programs or portfolios; evaluation of any “cut” or “slice” of the organization’s research (E.g. research on a particular theme, or in a particular geographic area etc.); or evaluation of all of the research that the organization conducts or funds.

Attachment 2
Participating Agencies and Interviews completed *

Agency / Organization	Number of interviewees by category			
	Agency	Program	Grantee	Evaluator
Swedish International Development Agency (Sida)	1		1*	[* also an Evaluator]
National Science and Technology Development Agency (Thailand)	1			
International Centre for Diarrhoeal Disease Research (Bangladesh)	1			
Gates Foundation (India)	1*			[* also an Evaluator]
United States Agency for International Development (USAID)	1			
Netherlands Foundation for the Advancement of Tropical Research (WOTRO)	1			
Global Health Research Initiative (Canada)	1	1		
International Development Research Centre (Canada)		1	2	1
Fiocruz (Brazil)	1			
National Research Foundation (South Africa)			1	
TOTALS	8	2	4	1 (+ 2)

** Note: two additional interviews with representatives of regional organizations were conducted during the June 2-3, 2014 workshop, bringing the total number of interviews to 17.*

Attachment 3 – Research Design

The study consists of two parts, with some overlap between them. Part 1 involves data collection and preliminary analysis, with Part 2 focusing on synthetic analysis. Summarized below are the research questions identified for each part, along with the methods that were used to address them.

Part 1: Descriptive data collection and preliminary analysis:

Research Questions:

1. *How are research funders and grantees evaluating global health research for development (GHR4D), with respect to approaches and frameworks?*
2. *What successes, best practices, challenges and solutions are being used to evaluate GHR4D at the organizational, program and project levels?*

Methods:

1. Update documents about EGHR4D:

Early discussions with the project commissioners indicated that we should focus on impact evaluation. We therefore searched for documents describing evaluations of health research, some part of which was being conducted in one or more LMICs. Documents had to have been published in the last ten years, and be available in English, Spanish or French. The searches themselves were however specified only in English.

The following topics were excluded in order to keep a tight focus and to restrict the volume of material to be analyzed to a manageable amount:

- Evaluation of health research capacity strengthening (HRCS). Health research systems in developing countries often lack capacity, so international development agencies commonly fund HRCS activities. We only included evaluations of HRCS when research was also being funded and there was a component of the evaluation assessing the impact of that funded research. As strengthened research capacity is one of the potential impacts of research, the indicators and metrics used in evaluations of HRCS are likely also to be of interest to evaluators of GHR4D. (Cole, *et al* 2014) ¹ provides a good starting point for readers who would like to explore this further. In addition, (Boyd, *et al* 2013) provides an analysis of frameworks used by development agencies to evaluate HRCS.
- Evaluation of knowledge translation and exchange to influence health policy and practice, unless it was in the context of evaluating a particular research project, program or product. Thus, research on health policy processes in LMICs, and the roles played by various

¹ References noted in this Attachment can be found in Attachment 7.

stakeholders, including their attitudes to research evidence, were excluded (E.g., (Cheung, *et al* 2011, Orem, *et al* 2013)).

- National systems for monitoring and assessing the research undertaken by individual researchers and higher education institutions (E.g., (Lange and Luescher 2003, Masipa 2011)).
- Research priority and agenda setting (E.g., (Landes, *et al* 2013)). See (Viergever, *et al* 2010) for a discussion of methods and good practices.

a. From the peer reviewed literature:

We had intended to search relevant databases (e.g. Medline/PubMed/Global Health), checking the publication title, abstract and keywords using a search string of the form (monitoring OR evaluation OR impact OR [list of evaluation methods and approaches]) AND research AND (global OR LMICs OR [list of particular LMICs, regions and continents]) AND health [for non-health databases only]. However, a pilot search of Medline produced 1500 “hits”, of which less than ten proved to be relevant. A similar search of a particular journal – Health Research Policy and Systems – found a much higher proportion of relevant documents, but predominantly only of background interest. We therefore proceeded by means of Google Scholar citation and similar articles searches, starting from documents that had already been found to be highly relevant and from key articles describing research evaluation frameworks and methods. Similar citation searches were used to identify these key articles.

In order to identify themes, descriptions of approaches and frameworks were displayed in matrices, along with background context information about how they were used and in what circumstances, and what the accompanying challenges were.

b. From the non-peer reviewed literature:

The starting point for the search was documents recommended to the project team by interviewees—particularly the “evaluators” – and key documents cited in peer-reviewed articles. Scans and searches of some websites were also conducted (E.g., [Overseas Development Institute \(ODI\)](#), [Institute of Development Studies \(IDS\)](#), [Active Learning Network for Accountability and Performance in Humanitarian Action \(ALNAP\)](#), [Itad, Research to Action](#)). It did not prove possible within the time available to search research funding agency websites. Similar approaches to text analysis (as described above) were used for this category of literature.

2. Collect primary and secondary data:

Early in the project, the study team and the funding advisory group discussed the issue of which funding agencies should be invited to contribute to the study. Criteria were agreed upon, and with further discussion, ten (10) agencies were identified.

[Note: as the project proceeded, some changes in this list were made, as a result of agency-specific considerations.]

Information from this source was obtained through semi-structured telephone or Skype interviews, as well as from surveys (see details below). Drafts of the interview guides and the survey instrument were helpfully reviewed by the funding advisory group. The interview guides were tailored to match the expected views of the three groups below— funding agency representatives, program managers and grantees. As the project proceeded, a fourth category emerged for inclusion, namely “evaluators”, both from within an agency as well as external evaluators.

a. From funding agencies:

Representatives of funding agencies were provided with background information about the project, and asked to participate in an interview and also recommend program officers and grantees to be contacted. As indicated above, some agencies also recommended evaluators.

b. From programs within funding agencies:

Where appropriate, agency representatives recommended program managers to be interviewed. As appropriate these individuals were provided with background project information as well as the interview guide (which listed questions for the interview).

c. From projects/grantees:

Researchers who had received project funding (that is, grantees) and who were recommended for inclusion in the study were invited to participate in a web-based survey. A few of these grantees were also interviewed, if they were judged to have some special interest in evaluation and had tested evaluation strategies.

d. From evaluators:

As indicated above, several individuals with expertise in evaluation were recommended by some agencies. Interviews with these individuals were very helpful. In several cases, they recommended additional documents and other information sources that enriched the study analysis.

e. From additional groups:

In addition to the above, we identified a number of networks/groups that are involved in evaluation methods and strategies. Representatives of these groups were also invited to complete the survey.

3. Conduct preliminary analysis:

- a. Thematic analysis of text material--both peer-reviewed and grey literature, was conducted using appropriate tools for qualitative analysis. Interview notes were analyzed in a similar fashion.

- b. Initially identified themes were reviewed by members of the project's implementation team (Neufeld, Cole, Boyd, Njelesani) for reliability and clarity. This was done through electronic communications and weekly teleconferences.

Part 2: Synthetic analysis:

Research questions:

3. *Are the identified approaches converging or diverging, and how?*
4. *What evaluation theories and frameworks are best suited to evaluate GHR4D?*

Methods:

- a. Building on earlier work by this study team on frameworks, and inventory of approaches, theories and frameworks was developed as a tool for depositing and updating data as the project progressed.
- b. A matrix analysis tool was used for cross-case comparisons across frameworks and approaches, to judge the extent of convergence and divergence, and other relevant comparisons.
- c. The team then put forward its judgments as to which evaluation theories and frameworks seemed best suited to GHR4D. These recommendations are accompanied by "good practice" examples from both the document analysis and the interviews.

Attachment 4a - Summary of findings from the literature review

Each of the documents included in our literature review is summarized in the table in Attachment 4b. The table is in two parts:

1. Evaluations of the impact of GHR4D, typically funded by research funders or grantees;
2. Systems and processes for evaluation within research funder organizations.

These documents and the table are potentially useful as a source of ideas for evaluation approaches, and they highlight various contextual factors that evaluators might be wise to take account of, such as: staff/organizational turnover in developing countries which might pose difficulties for longitudinal evaluations; the need for commitment of time, effort and appropriate expertise in order to maximize the benefits accruing from evaluation; the potential for bias in retrospective, self-assessment based evaluations; 'insider-outsider' evaluator dynamics; the potential difficulty in accessing high-level officials when trying to evaluate policy impact; the variable quality and coverage of routinely collected administrative data in LMICs; the long timescales for some impacts to appear; that the funding of research in LMICs by international organizations is not a politically neutral act, and so politics can constrain and bias evaluations.

Most documents describe evaluations, but do not reflect extensively on the quality of the evaluation process, beyond some consideration of methodological limitations and how the evaluators have tried to address them.

The documents included in the review concerned research on a range of topics, conducted in a range of countries, with African countries most heavily represented.

Informing policy/practice was the impact most often evaluated, but there were few attempts to evaluate health benefits or socio-economic benefits. Where some assessment of health benefits was made, this was not the main focus of the evaluation, and was limited to listing examples, with no attempt at quantification or assessing value for money. Evaluations focused on more easily measurable and shorter-term impacts: data on academic publications in international journals is readily available, for example, and can provide a variety of quantitative indicators of knowledge advancement. Assessing whether research is informing policy/practice is perhaps more challenging, but is of obvious interest to development agencies. These evaluations were typically based on stakeholder judgments about impacts and their attribution to research, whether gathered through interviews, surveys or document analysis, with some efforts to triangulate data from different sources in an attempt to counteract potential bias. Most of these evaluations focused on policy rather than on practice.

Conceptual frameworks of the process of transferring knowledge into policy/practice were commonly used. This was partly to take account of various contextual factors, but also because many of the evaluations were not purely, or

even mainly, impact evaluations, but emphasized process evaluation in order to learn about what works in achieving impact. This may be a reflection of the difficulty and resource-intensiveness of impact evaluation, and of the potential of process evaluations to facilitate useful learning and improvement.

Some of the evaluations attempted to assess the impact of a country or region's research (or research system), or of the research of several countries, using either comparative statistics or comparative case studies. Other evaluations assessed projects. There was a relative paucity of evaluations assessing programs.

Relatively well known general research evaluation methodologies such as the Payback model (Buxton and Hanney, 1996) were not explicitly used in most evaluations reviewed, although some were referenced. Those "named" approaches that were used were based on experiences in developing countries. This could indicate either appropriate use of tailored instruments and lack of awareness of relatively well-known methodologies.

The evaluations were generally relatively unsophisticated such as more formal methods for data collection and analysis, more use of theories of change type approaches. However some did use multiple methods and triangulation. The quantitative approaches did not reflect on the appropriateness of their sampling strategies and statistical tests of significance were seldom mentioned, let alone used. Most of the qualitative evaluations did not attempt to produce generalizable knowledge through use of theory-based evaluation approaches such as realist evaluation, using instead simple thematic analysis techniques. Most inquiries were cross-sectional and retrospective, with no comparison groups and few longitudinal designs, although there were some analyses of trends.

There was an absence both of participative research methods being evaluated and of use of participatory evaluation methods, despite their potential for supporting capacity strengthening and development. Stakeholder involvement in the design and conduct of evaluations was typically absent, with just a few instances of involvement in selecting data collection sites and in validating findings.

The literature review process

We searched for descriptions, evaluations or reviews of the use of methods, tools, approaches and frameworks to evaluate health research in low and middle-income countries (LMICs). In particular, we looked for evaluations of impacts of GHR4D. We did not formally assess the quality of the evaluations.

Excluded from this review were articles with the following features:

- they were not focused specifically on health research in LMICs;
- they focused on health research capacity strengthening activities rather than research activities per se;
- they dealt with knowledge transfer and exchange activities, rather than research;
- they did not include substantive evaluations;

- the concerned models of the research impact process and relevant contextual factors, but not research as such.

Attachment 4b – Documents reviewed

Documents assessing the impact of GHR4D: Characteristics, approaches used and points of interest identified

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Adam, T., Ahmad, S., Bigdeli, M., Ghaffar, A. & Rottingen, J.-A. (2011) Trends in Health Policy and Systems Research over the Past Decade: Still Too Little Capacity in Low-Income Countries. PLoS ONE, 6, e27263. http://dx.doi.org/10.1371/journal.pone.0027263	health policy and systems	advancing knowledge; HRCS	LMICs	LICs overall compared with MICs and HICs (data at research institution level provided by individual survey respondents)	Bibliometric analysis, repeated survey of research institutions	Example of a quantitative approach looking at trends (bibliometric analysis plus repeated cross-sectional surveys)
Casale, M. (2006) Bridging The Gap Between Researchers and Policymakers: HIV in South Africa. HEARD, University of KwaZulu-Natal, Durban. www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/3687.pdf	Child health and wellbeing	Informing policy/practice	South Africa	Project (impact of 1 dissemination conference on district-level stakeholders)	Outcome Mapping ODI RAPID framework Questionnaires, structured interviews, performance journal	Detailed description and analysis of the use and utility of Outcome Mapping in practice.

¹ Although not all articles were purely focused on impact assessment, they all contained at least some impact assessment.

² By this we mean the totality of the research about which the evaluation sought to make judgments (sometimes based on a sample of the research) This needs to be distinguished from the level at which the impacts of the research were assessed. For example, a single research project might have impacts at local, national, regional or global levels.

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Dagenais, C., Queuille, L. & Ridde, V. (2013) Evaluation of a knowledge transfer strategy from a user fee exemption program for vulnerable populations in Burkina Faso. Global Health Promotion, 20, 70-79. http://ped.sagepub.com/content/20/1_suppl/70.abstract	health services	advancing knowledge; informing policy/practice	Burkina Faso	Project	Quantitative questionnaire (claimed), document analysis, interviews with key informants	Categories of conditions that encourage research use (Table 1); Types and examples of knowledge use by different target groups (Table 2)
Djalalinia, S., Peykari, N., Owlia, P., Eftekhari, M.B., Habibi, E., Falahat, K., Mojarrab, S., Malekafzali, H. & Ghanei, M. (2013) The analysis of health research system evaluation in medical sciences universities. Iranian journal of public health, 42, 60-65. http://europepmc.org/abstract/MED/23865018	health	advancing knowledge; HRCS	Iran	National health research system	Modified version of WHO Health Research System Analysis (HRSA) indicators	Indicators/measures used; trend analysis
Druce, N., Gopalan, A., Moree, M., Nathanson, N., Plotkin, S. & Skolnik, R. (2009) Evaluation of the International AIDS Vaccine Initiative 2003-2007. The World Bank. www.iavi.org/Documents/World%20Bank%20Evaluation%202009.pdf	HIV/AIDS	Advancing knowledge; HRCS; informing policy/practice	India, Kenya, Uganda, South Africa (IAVI office sites)	Programme/Organization (1 funder - IAVI)	Stakeholder interviews, document analysis, secondary data analysis, field visits.	Different measures/concerns when research is early stage and more for product development than for publication

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Gholami, J., Ahghari, S., Motevalian, A., Yousefinejad, V., Moradi, G., Keshtkar, A., Alami, A., Mazloomzadeh, S., Vakili, M., Chaman, R., Salehi, B., Fazelzadeh, O. & Majdzadeh, R. (2013) Knowledge translation in Iranian universities: need for serious interventions. Health Research Policy and Systems, 11, 43. http://www.health-policy-systems.com/content/11/1/43	medical	informing policy/practice	Iran	National (based on 9 universities/research institutes)	Self-Assessment Tool for Research Institutes (SATORI), administered in focus groups	Use of focus groups alongside a structured assessment tool; themes derived from the tool data
Hamann, S., Mock, J., Hense, S., Charoenca, N. & Kungskulniti, N. (2012) Building tobacco control research in Thailand: meeting the need for innovative change in Asia. Health Research Policy and Systems, 10, 3. http://www.health-policy-systems.com/content/10/1/3	tobacco control	advancing knowledge	Thailand	National	Document analysis, interviews with key individuals, bibliometric analysis.	
hera (2014) End of Programme Summative Evaluation of the Africa Health Systems Initiative Support to African Research Partnerships (AHSI-RES). Volume 1: Main report. Volume 2: Annexes. The Global Health Research Initiative.	human resources for health and health information systems	HRCS, informing policy/practice	Burkina Faso, Kenya, Mali, Malawi, Tanzania, Uganda, Zambia	Programme (10 projects) (GHRI)	Tabulation of indicator results against targets Theory of change Document review, interviews, participant observation, survey	Comparisons with baseline data. Annexes contain copies of the data collection instruments used E.g. interview guides.

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Ismail, S.A., McDonald, A., Dubois, E., Aljohani, F.G., Coutts, A.P., Majeed, A. & Rawaf, S. (2013) Assessing the state of health research in the Eastern Mediterranean Region. Journal of the Royal Society of Medicine, 106, 224-233. http://jrs.sagepub.com/content/106/6/224.abstract	health	advancing knowledge; informing policy/practice	Countries in the Eastern Mediterranean	Region (based on assessment of National health research systems)	Rapid assessment of peer-reviewed and grey literature Systems approach to evaluation [claimed]	Indicators/measures identified; systems approach to evaluation
Kok, M.O., Rodrigues, A., Silva, A.P. & de Haan, S. (2011) The emergence and current performance of a health research system: lessons from Guinea Bissau. Health research policy and systems / BioMed Central, 10, 5. http://europepmc.org/abstract/ME D/22321566	health	HRCS; informing policy/practice; health benefits	Guinea Bissau	National health research system	Case study (document analysis, interviews of key informants) Conceptual framework: The National Health Research System (NHRS) framework (Table 1).	
Mahmoud, A.O., Ayanniyi, A.A., Lawal, A., Omolase, C.O., Ologunsua, Y. & Samaila, E. (2011) Perceptions of Nigerian medical specialists on research. Journal of Public Health in Africa, 2, e1.	medical	advancing knowledge; informing policy/practice	Nigeria	National level (sample of individual researchers in 6 tertiary health institutions)	Structured questionnaire survey of medical specialists	Example of a quantitative survey approach; indicators/measures used

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Majdzadeh, R., Rahmani, K. & Nasehi, M. (2013) What Is the Share of the Country's Researches in Iran's National Tuberculosis Guideline? Iranian Journal of Public Health, 42, 1405-1413. http://ijph.tums.ac.ir/index.php/IJPH/article/view/4713	TB	informing policy/practice	Iran	National health system	Appraisal of clinical guideline (literature review, interviews) Oxford Centre for Evidence-based Medicine tool for determining levels of evidence	Appraisal of a clinical guideline as a way of assessing national research impact
Mirzoev, T.N., Omar, M.A., Green, A.T., Bird, P.K., Lund, C., Ofori-Atta, A. & Doku, V. (2012) Research-policy partnerships - experiences of the Mental Health and Poverty Project in Ghana, South Africa, Uganda and Zambia. Health research policy and systems / BioMed Central, 10, 30. http://europepmc.org/abstract/MED/22978604	mental health	advancing knowledge	Ghana; South Africa; Uganda; Zambia	Project (research-policy partnerships between 5 research institutions (within country) and 4 Ministries of Health)	Semi-structured interviews with respondents from the research teams and Ministry of Health partners Conceptual framework (figure 1)	Conceptual framework; potential negative effects of research-policy partnerships; triangulation of information to minimise bias; value of 'insider-outsider' perspectives.
NWO-WOTRO Monitoring & evaluation plan for the Dutch Global Health Policy and Health Systems research programme. NWO-WOTRO MTR (2005-2008) form: testable goals (review questions). NWO-WOTRO.	Global health; health systems	Advancing knowledge; HRCS; informing policy/practice	Not specified	Programme and project (NWO-WOTRO)	Mid-term (self-assessment) and final (external) reviews Stories of most significant change, logframe Cohort analysis of a cohort of research projects after 10 years	The mid-term review form contains some examples of measures. Plans/forms only – no empirical data.

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Rispel, L.C. & Doherty, J. (2011) Research in support of health systems transformation in South Africa: The experience of the Centre for Health Policy. J Public Health Pol, 32, S10-S29. http://dx.doi.org/10.1057/jphp.2011.33	health policy	advancing knowledge; informing policy/practice; HRCS	South Africa	Research centre within a University	Document reviews; semi-structured key informant interviews	Triangulation of information to minimize bias; value of 'insider-outsider' perspectives.
Sridharan, S., Maplazi, J., Vijendran, M., Richardson, E., Nichols, J. & Parekh, H. (2013) Final report : evaluation of Teasdale-Corti Initiative. Evaluation Centre for Complex Health Interventions (TECCHI), St. Michael's Hospital. http://hdl.handle.net/10625/52321	Health problems	Advancing knowledge, HRCS, informing policy/practice	LMICs, including Chile, China, Honduras, Mexico, Nigeria, South Africa, Sri Lanka, Thailand	Programme (14 projects) (GHRI)	Interviews, document analysis, surveys, interviews, bibliometric analyses, brief case studies.	Shows how theory of change-based evaluation and impact evaluation can be combined. Incorporates analysis of both research proposals and reports. Identifies issues regarding the design of monitoring and evaluation systems
Ssengooba, F., Atuyambe, L., Kiwanuka, S., Puvanachandra, P., Glass, N. & Hyder, A. (2011) Research translation to inform national health policies: learning from multiple perspectives in Uganda. BMC International Health and Human Rights, 11, S13. http://www.biomedcentral.com/1472-698X/11/S1/S13	HIV-prevention	informing policy/practice	Uganda	Project (2 projects in one university)	Case studies (in-depth interviews with key stakeholders) Various conceptual frameworks linking the research-policy interface.	Conceptual frameworks; Preferred characteristics and sources of evidence for policy in Uganda (table 4).

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Theobald, S., Taegtmeier, M., Squire, S., Crichton, J., Simwaka, B., Thomson, R., Makwiza, I., Tolhurst, R., Martineau, T. & Bates, I. (2009) Towards building equitable health systems in Sub-Saharan Africa: lessons from case studies on operational research. Health Research Policy and Systems, 7, 26. http://www.health-policy-systems.com/content/7/1/26	health	informing policy/practice	Kenya; Malawi; Nigeria	Project (comparison of 3 projects - aiming for theoretical generalization)	Case studies ODI RAPID framework for research-policy links (Figure 1)	Sampling for theoretical generalization (as opposed to statistical generalization); conceptual framework
Tulloch, O., Mayaud, P., Adu-Sarkodie, Y., Opoku, B., Lithur, N., Sickle, E., Delany-Moretlwe, S., Wambura, M., Chagalucha, J. & Theobald, S. (2011) Using research to influence sexual and reproductive health practice and implementation in Sub-Saharan Africa: a case-study analysis. Health Research Policy and Systems, 9, S10. http://www.biomedcentral.com/1478-4505/9/S1/S10	sexual and reproductive health and HIV	informing policy/practice	Ghana; South Africa; Tanzania	Project (3) / Program (1)	Multiple case-studies (critical reflection by researchers involved) Conceptual frameworks: Sumner's framework for understanding the process of transferring research evidence into policy/practice; Walt and Gilson's policy triangle; research use continuum - Nutley et al	Use of criteria for purposive sampling; conceptual frameworks

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
Whiteside, A. & Henry, F. (2011) The impact of HIV and AIDS research: a case study from Swaziland. Health Research Policy and Systems, 9, S9 http://www.biomedcentral.com/1478-4505/9/S1/S9	HIV/AIDS	informing policy/practice	Swaziland	Project (impact of 1 research report on national-level and relevant international stakeholders)	Literature review, document analysis, questionnaire, interviews	Pragmatic, practical approach to assess the impact of a research report; attempts to address bias
Documents focusing on systems and processes for evaluation within research funder organizations:						
Carden, F. & Earl, S. (2007) Infusing evaluative thinking as process use: The case of the International Development Research Centre (IDRC). New Directions for Evaluation, 2007, 61-73. http://dx.doi.org/10.1002/ev.243	health	Not specified	Not specified	Project (plus cross-cutting trends - IDRC organization)	interview-based rolling Project Completion Reports at three stages of the project Annual Learning Forum	Tools and process for learning across a funder organization and from grantees that has demonstrated benefits
Cousins, J.B., Bourgeois, I. & Associates (2014) Multiple Case Study Methods and Findings. New Directions for Evaluation, 2014, 25-99. http://dx.doi.org/10.1002/ev.20077	health	Not specified	Not specified	Project (plus cross-cutting trends - IDRC organization)	interview-based rolling Project Completion Reports at three stages of the project Annual Learning Forum	Tools and process for learning across a funder organization and from grantees that has demonstrated benefits

Document	Research Topic/ Issue	Focus of impact assessment ¹	LMICs where research situated	Level at which research assessed [see Note ²]	Approaches used	Points of interest identified
MacLean, R. (2012) The Global Health Research Initiative: A preliminary review and analysis of potential cross-program impacts. unpublished report. Canadian Institutes of Health Research. Institute of Population and Public Health.	Global health	Advancing knowledge, HRCS, informing policy/practice, health benefits, socio-economic benefits	Not specified	Organization/ Cross-programme (5 programmes) (GHRI)	CAHS impact framework Document review	Lists CAHS framework categories of indicators for each impact, with some specific examples of indicators. Identifies various reporting issues if programmes are to be compared.
Seifried, A. (2013) Taking stock of GHRI evaluations and reviews: what we have, what we need and suggestions for the future evaluation plan. First draft. Unpublished report. GHRI-IRSM.	Global health	HRCS, informing policy/practice, health benefits	Not specified	Organization ("platform") (GHRI)	Document review of previous evaluations and reviews	Evaluation of evaluations Identifies issues to be addressed if an evaluation is to assess the added value that an organization brings to the research it funds.
Wind, T. & Carden, F. (2010) Strategy evaluation: Experience at the International Development Research Centre. New Directions for Evaluation, 2010, 29-46. http://dx.doi.org/10.1002/ev.344	health	Not specified	Not specified	Program/Organization (1 funder - IDRC)	Strategic evaluations to address cross-cutting issues and elements of corporate strategy.	Addresses strategic decision-making by a funder organization through an approach that has demonstrated benefits.

Attachment 5a – Survey Findings: a summary

1. 48 complete responses and 54 partial responses were received. Unless stated otherwise, our analyses are based on both complete and partial responses. The relatively low response rate (we know that at least 400 people were invited to take the survey) may indicate the relatively specialist nature of the topic and perhaps a lack of attention paid to it. The relatively high proportion of partial responses may reflect the difficulty of communicating the essence of this specialist topic in our invitation letter (E.g. one partial respondent emailed me to say that he is an experienced evaluator, but realised part way through the survey that we are only interested in evaluation of GHR4D and his experience is of evaluating research for development that is not health specific).
2. Respondents were generally in quite high-ranking positions in their organisations (Directors and Executives were common).
3. Respondents were from a mix of organisations (Universities in LMICs and in HICs; research funders; and development agencies/organisations). The majority of these organisations were research organisations rather than evaluation organisations (see Table 1).
4. Similarly, most respondents were researchers rather than evaluators (see Table 3). This might be indicative of a lack of evaluators specialising in evaluation of GHR4D, or a lack of evaluation of GHR4D (these would likely be related), or that our survey failed to reach such evaluators/get them to respond.
5. Partial respondents: a higher proportion specified that their organisations had other roles, and a lower proportion that their organisations funded or commissioned GHR4D. A higher proportion had roles which included evaluating GHR4D. A higher proportion indicated that stakeholders were conducting a lot of evaluation.
6. It would appear that there is scope for more evaluation to be done (see Table 4), particularly by governments, although neither research funders nor grantees would be characterised as doing a lot of evaluation of GHR4D. NGOs were mentioned by five respondents as another stakeholder that may conduct some evaluation. Even though respondents to this survey were likely to be more interested in evaluation than the norm, they did not characterise themselves as doing lots of evaluation either, although this may simply reflect that their main role is not as an evaluator, but as a researcher etc.. Resources for evaluation, commitment to evaluation, and more analytical evaluation were mentioned in several comments as being desirable, but there were divergent views regarding the value of process evaluation. Some comments also suggest that some respondents were not distinguishing between the evaluation of research, the evaluation of development projects and research itself (CF point 1 above). Some selected comments:
 - a. *“No real commitment of the government or the University in this field”*
 - b. *“Over-structured, pro-forma evaluations (the same for every project, country and care-group) can hardly be very informative, especially when they focus on research activity rather than on what knowledge, capacity, policy etc. was produced”*
 - c. *“What might improve the situation? Building evaluation into all funded projects
Funding to carry out evaluation activities
Increasing duration of funded projects to enable evaluation”*
 - d. *“Time. The orientation around evaluation must move past impact, towards more qualified, nuanced understandings of research evaluation, such as realist evaluation and other theory-driven approaches, which are necessarily data, time and resource heavy, which often runs counter to granting time cycles”*

- e. *“More active participation by stakeholders in LMICs”*
 - f. *“Developing a culture of evaluation and making it part of a general implementation strategy”*
7. The funder and research organisations represented do appear to have a practical/development focus rather than a narrow academic focus (see Table 2) – the aim of advancing knowledge is no more common than the aim of producing health benefits. This is reassuring in terms of accessing the right group of respondents. The relatively high emphasis on producing health benefits may indicate that more evaluations should be trying to evaluate health benefits (CF the analysis of peer-reviewed articles, where few if any evaluations of the evaluations found were trying to do this). However, comparing Table 5 with Table 4 suggests that overall the impacts that evaluations seek to assess are in proportion to the impacts that GHR4D is seeking to achieve. That evaluations of health and socio-economic benefits are not more prominent in the peer-reviewed literature might indicate that such evaluations currently lack rigour/robust findings [CF barriers to impact evaluation highlighted in the intro], and that there is a need to tackle these barriers. E.g., one respondent said *“We have tried to look at wider socioeconomic issues- but too costly for us....all done on a shoe string”*.
8. Various suggestions were made about how to help ensure that the “right” impacts are evaluated, including better systems for routine collection of health data, capacity building and stakeholder involvement/accountability to stakeholders. Selected comments:
- a. *“Clear statements of objectives and indicators to measure them helps stakeholders, particularly if the are agreed upon by all stakeholders”*
 - b. *“Health impacts are notoriously difficult to assess. Experience from the Tanzania Essential Health Interventions Project (TEHIP) with which I was involved for many years pointed to the real need to have a system in place to track mortality in large populations as well as to assign causes of death.....through demographic sentinel surveillance systems (DSS).”*
 - c. *“Right impacts can be assessed by the help of collection of secondary data. But the response of concerned dept in giving the required data is more important.”*
 - d. *“This mean that getting all procedure transparent, consider the value of accountability, feed-back to other stakeholders about the project progress and potentials challenges is very helpful.”*
 - e. *“the selection of these impacts should be influenced by what decision-makers and practitioners would perceive to be compelling impact areas. What doesn't help is to generate impact statements that are disconnected with the intended users of the findings. Greater interaction with intended users throughout the evaluation process can help avoid this mismatch.”*
 - f. *“We need a better evaluation approach. I am tired of "frameworks" and fancy jargon.”*
 - g. *“Consistent, constant evaluation embedded, e.g. on health systems (as opposed to isolated evaluation exercises). Informed policy and practice mechanisms can (and should) include evaluative operational research. One-time only evaluation interventions can be deceiving.”*
9. Survey responses suggest that there is less emphasis on evaluating research at a whole organization level than there is on evaluating projects and programmes (Table 6). One respondent highlighted leadership change (and hence policy change) within research funders and governments as a barrier to evaluation. This inference must be treated with

caution however, as the sample is small and there is scope for misinterpretation of the question. Selected comments:

- a. *"We evaluate program success both at the project level and overall. In my particular capacity in my agency, I also am seeking to evaluate our overall research investments (multiple programs). This is not something that my agency routinely does however"*
- b. *"For research funders, evaluations are often to evaluate the program to determine if this is a good approach, and to inform programmatic decisions. Less frequently, the institution commissions an evaluation at the organizational level to assess its overall programmatic approach, or that within a certain thematic area."*

10. Survey responses indicate that there is scope for improving the coordination of evaluation practices between GHR4D funders, grantees and other stakeholders (Table 7), and that this would enhance evaluations. Selected comments:

- a. *"Coordination is development work has generally been a challenge because of organisational differences, leadership and other factors. Coordination is effective only when there is scramble for perceived benefits"*
- b. *"Companies and health academics need to work more closely to develop innovations and evaluate them."*
- c. *"A more consultative, flexible approach by the funder, which is not driven by the need to "show results". At times there are no results, i.e., the evaluation is 'empty', and this needs to be acceptable to funders as well. Timelines need to be longer, and evaluative accountabilities need to be multi-polar, not simply oriented to the funder in order to capture all impacts and processes of whatever the intervention has introduced"*
- d. *"There are some structural elements that can aid in coordination such as coalitions and consortia of different research funders, research organizations and decision-makers. These types of bodies can help convene groups and provide space to share what they have accomplished (what most people want to talk about) and also what people are planning to do (where the most interesting potential coordination can take place)."*
- e. *"Even when co-ordinated - those on ground not asked what they would like to see measured or what they see as success....the metrics are all very "come from away" metrics We in the west often do not even realize when we shut down input from those on the ground - and they do not want to "bite the hand that feeds them ""*
- f. *"National health research agenda. National research coordination mechanisms. National body to provide oversight and regulation"*
- g. *"Big researchers events (symposia, fora, etc) are not very helpful to aid coordination, at least they seldom have coordination as a common objective. Major funding agencies or international organizations (such as WHO) can help establish research and evaluation agendas. Might help: a mechanism bringing together major funders and commissioners to identify demand, opportunities, and methodologies for evaluation"*

11. Experiences of using particular approaches, frameworks or methods to evaluate global health research for development. Selected comments:

- a. *"In our experience, we have found that evaluating GHR4D always requires a mixed methods approach, often using a developmental evaluation lens. Outcome mapping and outcome harvesting approaches have worked well."*
- b. *"It's important to specify a program pathway or theory of change as a framework to understand the process of change and impact."*

- c. *"I have had good experiences using the utilization-focused evaluation approach. This is less of a framework or method and more of an overarching philosophy when designing and implementing an evaluation study. I have worked with research networks supporting global health research in Asia and Africa. Initially it is a bit difficult for people to step away from particular instruments, methods and frameworks (as this is how most people are trained) and to think more broadly about how the evaluation is focused. However, over time and ongoing dialogue, it is clear that the researchers grow more critical and take more ownership of the evaluation designs, practices and related findings that emerge."*
 - d. *"Usual process metrics - so can reassure funders- number participants trained, projects developed , completed, published etc Then asking open ended questions, providing opportunities to really hear form locals impacted by work - thru FGD, surveys, online networks etc. Not just asking the usual - participants, government , local leaders etc but also HCW in field not in academe, patients etc. Working with locals to set framework for other evaluations that fits their goals and outcomes for the program NOT our goals Re-checking that what measuring is what is valued not just what is usually counted Also being open to seeing unintended consequences - both good and bad - and collecting that information. Going back after ++ time (several years) to see if sustained change Trying to assess how much they own the program versus they participate in program We have yet to find a really good evaluation framework that does all of these things.....so many are process focused"*
 - e. *"Attempts by WHO to evaluate all of its own research efforts (commissioning, proposing agendas, funding) have failed, because of excessive "departmentalization". An informal governance that promotes regular meetings among departments that dedicate mostly to research promotion has been more effective in coordinating and promoting evaluation of research activities."*
12. Only 30% of respondents offered to provide access to documents they thought might be relevant. This may be indicative of a lack of documents to support EGHR4D, or that the documents that exist are not accessible. Only a few respondents supplied URLs of documents they thought might be relevant, and these are not particularly focused on EGHR4D:
- a. Outcome Mapping - <http://www.outcomemapping.ca/> - possibly relevant as a general methodology
 - b. Tackling the tensions in evaluating capacity strengthening for health research in low- and middle-income countries - <http://heapol.oxfordjournals.org/content/early/2014/04/08/heapol.czu016.full> - as one of the authors, this is gratifying!
 - c. My M&E - <http://www.mymande.org/> - various reports and checklists, but not focused on EGHR4D.
 - d. Rand - http://www.rand.org/pubs/technical_reports/TR629.html - this compares 8 health research evaluation frameworks – so relevant background.

Table 1: Roles of respondents' organisations in relation to global health research for development (multiple answers permitted) [Complete responses only]

Answer	Response	%
Doing GHR4D	51	62%
Funding or commissioning GHR4D	9	11%
Evaluating GHR4D	18	22%
Other roles	26	32%

Table 2: What respondents' organisations seek to achieve through conducting or commissioning GHR4D (multiple answers permitted) [Complete responses only]

Answer	Response	%
Advancing knowledge	39	80%
Strengthening research capacity in developing countries	37	76%
Informing decision making by health policymakers or practitioners in developing countries	39	80%
Producing health benefits in developing countries	36	73%
Producing wider socio-economic benefits in developing countries	22	45%
Other aims	8	16%

Table 3: Roles of respondents in relation to global health research for development (multiple answers permitted)

Answer	Response	%
Doing GHR4D	38	58%
Evaluating GHR4D	12	18%
Building capacity to conduct GHR4D	34	52%
Building evaluation capacity/enabling others to evaluate GHR4D	16	24%
Other roles	9	14%
Funding or commissioning GHR4D	7	11%

Table 4: The amount of evaluation of GHRD commissioned or conducted by key stakeholders

Question	Little or no evaluation	Some evaluation	A lot of evaluation	Total Responses
Research funders	28%	48%	24%	54
Grantees / Research organizations	27%	53%	20%	55
National or regional governments in LMICs	71%	21%	8%	52
Other stakeholders	31%	62%	7%	13
National or regional governments in HICs	55%	31%	14%	51
Yourself	30%	53%	18%	40






Table 5: The impacts of GHR4D that different stakeholders seek to assess through evaluation (multiple responses permitted)

Question	New knowledge	Strengthened research capacity	Informed policy and practice	Health benefits	Wider socio-economic benefits	Total Responses
Research funders	17	16	17	17	11	78
Grantees / Research organizations	16	14	20	20	14	84
National or regional governments in LMICs	2	2	3	6	6	19
Other stakeholders	5	4	6	6	5	26
National or regional governments in HICs	8	7	12	16	11	54
Yourself	13	16	13	10	9	61

Table 6: The organisational levels at which stakeholders seek to assess the value of GHR4D (multiple responses permitted)

Question	Individual GHR4D project level	Whole programme level (overall value of several GHR4D projects)	Whole organisation level (overall value of all its GHR4D)	Total Responses
Research funders	18	16	11	45
Grantees / Research organizations	16	11	9	36
National or regional governments in LMICs	5	4	2	11
Other stakeholders	5	4	3	12
National or regional governments in HICs	12	10	7	29
Yourself	14	10	7	31

Table 7: How the evaluation practices of GHR4D funders, grantees and other stakeholders are coordinated

Answer		Response	%
They typically consult each other before any one of them commissions or conducts an evaluation, so that they can identify evaluation questions, procedures and roles that will be useful to all stakeholders		2	5%
They typically inform each other before any of them commissions or conducts an evaluation, but decisions are made solely by the organization commissioning or conducting the evaluation		6	14%
There is typically little or no coordination		15	35%
The extent of coordination varies widely depending on which funders, grantees and other stakeholders are relevant to the research		19	44%
In other ways		1	2%
Total		43	100%

Default Question Block

Approaches, Frameworks and Methods for Evaluating Global Health Research for Development (GHR4D)

- This survey is part of a [project to identify best practices for evaluating GHR4D](#) which would be valuable for research funding agencies, researchers and other organizations. The project is led by the [Canadian Coalition for Global Health Research \(CCGHR\)](#), and includes researchers from Manchester Business School. It has been commissioned by the [Global Health Research Initiative \(GHRI\)](#), in collaboration with [IDRC's Evaluation division](#) and [ESSENCE on Health Research](#).
- GHR4D is the production of knowledge with the aim of understanding health challenges in low and middle income countries, and mounting an improved response to them. Such research may include measuring the problem; understanding its causes; elaborating solutions; translating the solutions or evidence into policy, practice and products; and evaluating the effectiveness of solutions. GHR4D reflects a commitment to health equity, and may have features such as long-term sustainable North-South partnerships, inter-disciplinary responses to complex issues, grounding in local contexts, and an orientation to policy or practice impacts.
- [More information about what we mean by evaluation of GHR4D](#)
- The **aim of the survey** is to gather information about how GHR4D is currently evaluated, and the approaches, frameworks or methods that have been used.
- The survey should take only 10-15 minutes to complete, and your responses will be confidential
- Use the [>>] and [<<] buttons at the foot of each page to move backwards and forwards through the survey.
- If you would prefer to complete the survey by email, or have any other queries, please email alan.boyd@mbs.ac.uk

Background

Information about you and your organisation to help us assess the survey coverage and interpret the results.

What is your job title?

What is the name of your organization?

What is **your organisation's role** in relation to global health research for development?
(Select all that apply)

- Doing GHR4D
- Funding or commissioning GHR4D
- Evaluating GHR4D
- Other roles (*please state*)

Which of the following aims does your organisation actively seek to achieve through /SelectedChoicesTextEntry
(Select all that apply)

- Advancing knowledge
- Strengthening research capacity in developing countries
- Informing decision making by health policymakers or practitioners in developing countries
- Producing health benefits in developing countries
- Producing wider socio-economic benefits in developing countries
- Other aims (please state)

What is **your** role in relation to global health research for development?
(Select all that apply)

- Doing GHR4D
- Funding or commissioning GHR4D
- Evaluating GHR4D
- Building capacity to conduct GHR4D
- Building evaluation capacity/enabling others to evaluate GHR4D
- Other roles (please state)

In your experience, how much substantial evaluation (as opposed to basic monitoring) of global health research for development do each of the following stakeholders commission or conduct?
(select one only for each stakeholder)

	Don't know	Little or no evaluation	Some evaluation	A lot of evaluation
Yourself	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Research funders	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grantees / Research organisations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
National or regional governments in LMICs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
National or regional governments in HICs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other stakeholders (please state)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please provide further details. What helps you, or other stakeholders, to conduct and commission the right amount of evaluation? What doesn't help? What might improve the situation?

For those stakeholders that do more than just a little evaluation, **which impacts of GHR4D do their evaluations seek to assess?**

(select as many as apply for each stakeholder)

	Don't know	New knowledge	Strengthened research capacity	Informed policy and practice	Health benefits	Wider socio-economic benefits
Yourself	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Research funders	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Grantees / Research organisations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
National or regional governments in LMICs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
National or regional governments in HICs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$(q://QID54/ChoiceTextEntryValue/4}	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide further details. What helps you, or other stakeholders, to assess the right impacts? What doesn't help? What might improve the situation?

For those stakeholders that do more than just a little evaluation, **at which levels do their evaluations seek to assess the value of GHR4D?**

(select as many as apply for each stakeholder)

	Don't know	Individual GHR4D project level	Whole programme level (overall value of several GHR4D projects)	Whole organisation level (overall value of all its GHR4D)
Yourself	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Research funders	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Grantees / Research organisations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
National or regional governments in LMICs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
National or regional governments in HICs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$(q://QID54/ChoiceTextEntryValue/4}	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide further details. What helps you, or other stakeholders, to focus evaluations on the level(s) at which they will produce the most benefit? What doesn't help? What might improve the situation?

In your experience, how are the evaluation practices of GHR4D funders, grantees and other stakeholders coordinated?

(select one only)

- They **typically consult each other** before any one of them commissions or conducts an evaluation, so that they can identify evaluation questions, procedures and roles that will be useful to all stakeholders
- They **typically inform each other** before any of them commissions or conducts an evaluation, but decisions are made solely by the organisation commissioning or conducting the evaluation
- The **extent of coordination varies widely** depending on which funders, grantees and other stakeholders are relevant to the research
- There is **typically little or no coordination**
- In other ways (please specify)
- Don't know

Please provide further details. What has worked well in aiding coordination? What hasn't worked well? What might help improve the situation?

Block 2

Approaches, Frameworks and Methods for Evaluating Global Health Research for Development

Please tell us about your experiences of using particular approaches, frameworks or methods to evaluate global health research for development. What has worked well? What hasn't worked well?

Can you provide us with any information about documents or webpages which describe approaches, frameworks or methods that have been used to evaluate GHR4D?

These might include evaluation reports, case studies or examples of good practice.

(If you are not sure whether the documents are relevant, answer "Yes")

- Yes
- No

How would you like to get this information to us?

(Select all that apply)

- I will provide you with URLs/Web addresses for the information now
- I will email the information to you after completing this survey
- I would like you to give me a short phonecall so that I can provide you with a particular piece of information that is not in a public document (E.g. a good practice example)
- I would like you to give me a longer phonecall because there is much I would like to share with you about evaluating GHR4D
- I will give you the contact details of someone else who can provide you with the information
- In another way (please state)

Please list the documents/webpages and their URLs/addresses that you can provide us with now

	Document/Webpage name	URL/Web address
Doc 1	<input type="text"/>	<input type="text"/>
Doc 2	<input type="text"/>	<input type="text"/>
Doc 3	<input type="text"/>	<input type="text"/>
Doc 4	<input type="text"/>	<input type="text"/>
Doc 5	<input type="text"/>	<input type="text"/>
Doc 6	<input type="text"/>	<input type="text"/>
Doc 7	<input type="text"/>	<input type="text"/>
Doc 8	<input type="text"/>	<input type="text"/>

What is your email address, so that we can contact you to make arrangements for phoning/emailing information?

Email

Please provide the contact details of a person who can provide us with information about approaches, frameworks or methods for evaluating GHR4D.

Name

Email

Job title *(optional)*

Organisation *(optional)*

Telephone *(optional)*

Block 3

Networking

Please use the box below to tell us about other individuals or networks of people who you think might be interested in this project.

We would also be very happy for you to contact people directly to pass on the link to this survey

[\[http://mbs.az1.qualtrics.com/SE/?SID=SV_1Byx5QNYQaHstpz\]](http://mbs.az1.qualtrics.com/SE/?SID=SV_1Byx5QNYQaHstpz) or our project website [\[http://www.ccghr.ca/questionnaire-evaluating-global-health-research-development/\]](http://www.ccghr.ca/questionnaire-evaluating-global-health-research-development/)

The project commissioners are planning to use the project findings to develop a good practice document to help improve the evaluation of global health research for development.

If you would like to be notified about the publication of the good practice document, please type your email address here.

Email

By clicking on the [>>] button at the foot of this page your responses will be saved and passed to the project team.

Are there any comments you would like to make about evaluation of global health research for development before you finish the survey?

Attachment 6a – Summary of Interview Findings

What emerges from the responses of participants, irrespective of level i.e. agency, program, grant/project, is the complexities associated with evaluation particularly around defining and measuring impact (scientific and societal). While not particularly novel as an issue that most agencies and individuals involved in evaluation are faced with, it is worth noting that organizations, irrespective of level are still grappling with impact (i.e. attribution vs. contribution) suggesting there remains a gap within the evaluation sector. As one participant noted in responding to a question on the key results their agency seeks in global health research for development, measuring impact is challenging particularly with regard to societal health impact as they are unable to *actually trace down to health status change although the agency is interested in fewer maternal deaths*. A further challenge noted

Good Practice Taking Account of Research Funding Impact

WOTRO has adopted *Cohort Analysis* and incorporated it into its evaluation practices. This involves tracking previous recipients of research funding 5-10 yrs after the end of the funding period and requesting information on results of research (e.g. publications and any notable research outcomes and outputs) post funding period.

Through this analysis, WOTRO has observed that project outputs are realized mainly after the project period and particularly not within the first year after project ended with key publications occurring about 4-5 years after project funding had stopped. In one instance, cohort analysis revealed that a study with a very narrow focus on health research had considerable development relevance even though during the funding period the project had little relevance and connection to broader development. The cohort analyzed was funded 15 years prior. *We noticed that almost all the research we funded at the end turned out to be quite relevant for development even though the research was not assessed for development relevance during the project period.*

particularly by research funders was how to take account of research impacts that typically are realized long after the research-funding period has passed. This sentiment was captured by a participant who observed: *when research funding stops we don't have the means to continue measuring outcomes and impacts*. Longer time periods are seen as essential to fully understanding the impact and influence of research interventions while at the same time evaluation is critical within concurrent monitoring systems.

The majority of participants spoke of the need for a broader perspective of impacts beyond the funding period and societal wide and pointing to a need to move away from prioritization of publications as evidence of impact. These complexities extend to the perceived challenge of not having an

appropriate level of resources (financial & human) for evaluation of global health research; limited collaboration and sharing of evaluation approaches and experiences among donors and grantees (some grantees have never seen an evaluation of their donors); and consequently a diminishing role for global health research for development as a key component in development and global health planning as emphasized in key policy documents and programs in governments and

organizations. A notable challenge identified was that of capacity of grantees/recipients to negotiate the evaluation.

These complexities were readily identified by all participants regardless of level. However, there was no consensus on an evaluation approach or approaches that are ideal for the evaluation of global health research for development. Participants recognize the need to pay attention to context and need for flexibility in whatever evaluation approach is adopted as *in the past you had designs that required fidelity, now you need experimentation and flexibility*.

Within global health research for development there is an interesting dynamic occurring with the changing role of national research funding agencies with respect to the level of local resources (financial and systems) available for national health research in upper-middle income countries. These funding agencies are increasingly able to commission and support national health research with less support from traditional donors and research funders. At the same time these agencies are forging partnerships with other similar research funding agencies from upper middle-income countries. As these funding agencies grow, they increasingly have to address the challenges that come with a large bureaucracy while ensuring their evaluation and systems of accountability are comparable with traditional donors. Participants provided interesting comparisons of the rigors of evaluation for accountability to national governments and international donors.

Interestingly and perhaps unexpected, participants highlighted some of the geopolitical considerations associated with attaining upper-middle income status and distinction between national and global health research and the associated implications. Specifically when does global health research become national health research (and vice-versa) and what are the implications for a country being a recipient of global health research for development aid with regards to its perceived international standing as a middle-income country? This is outside the scope of the present study but important nonetheless to point out and consider as countries move towards higher income levels that potentially enable them to better address national (and/or global) health priorities.

Related to the geopolitical considerations is the observation of the growing role of regional health forums as mechanisms to support collaboration and knowledge sharing. With their political mandates these forums can be avenues to reinforce norms and priorities (health systems) within GHR. A key strength of some of these forums are their self-sufficiency i.e. funded by regional governments and not donor dependent and emphasis on context (regional) specific priorities. Whether its Africa, Asia or Latin America, ministries of health and other national health agencies are participating in many regional health forums that prioritize issues and practices such as the use of evaluation for decision making in health systems.

Participants at all levels point to a shift in global health research evaluation thinking from traditional evaluation approaches and designs emphasizing research needs

and capacities in low-income countries to a growing recognition of global challenges that affect all countries (e.g. emerging infectious diseases) and a need to better understand how evaluation can improve global health interventions as noted by a participant that spoke of *a shift from proving to improving global health interventions and programs*. Participants at the grantee and project levels see a greater role for themselves in informing and shaping policy and decision-making. Even with this shift however, research publications are largely considered the standard in evidence for research.

Evaluation Purpose:

Responses of participants regarding the purposes of evaluation can be grouped into the following five themes:

1. Accountability to Funders
2. Advancing Knowledge
3. Capacity Building
4. Informing Decision Making
5. Health and Social Benefits

Accountability and Advancing Knowledge (learning) and were predominantly cited as the main purposes of evaluation by all participants irrespective of level. Evaluation is seen as having a key role in demonstrating to funders/donors how resources were utilized and the results achieved, and also for program modification and advancing the field.

At the agency level, evaluation was also viewed as central to an agency's ability to influence decision-making and global health policy although participants could not offer specific ways in which decisions or policies could be influenced particularly at the global level. Evaluations were also viewed helpful to informing whether partnerships were worthwhile. Grantees noted that evaluation was particularly helpful in reflecting on strategies used to implement activities as well as understanding the characteristics of interventions that enable programs to grow and achieve desired objectives. The following response by a participant to the question on the purpose of evaluation is illustrative of this:

To be able to tell or show funders how money was used....an important element is learning for ourselves [at strategic level] what approaches work well and what does not. But also for our project, we want to keep an open mind and reflect on their approaches

Capacity building did not feature prominently although one program did have a capacity building in the health sector as an overall goal. Indeed most agency representatives noted the lower priority of capacity building as echoed in the following response by a participant.

A much lower focus is capacity building. It is something that I'm interested in but I would not say that it's an institutional priority.

What equally stands out with regards to evaluation is a lack of emphasis on evaluation capacity. All participants discussed capacity building with emphasis on the specific technical skills required for their particular research intervention. Capacity building in evaluation, based on responses of all participants is not a priority for many organizations although strengthened capacity in evaluation (institutional, individual) is seen as beneficial to the realization of the agency/program/grantee mission and objectives.

Closely related to strengthened capacity of an institution and individual was the ability of a grantee to negotiate the evaluation process such that their needs are prioritized throughout the evaluation process. This is particularly so for donor driven evaluations which most grantees frequently described.

Evidence Used In Evaluation:

The discussion on evidence was indicative of the shift participants noted is happening in evaluation and global health research for development. Specifically less emphasis on bibliometric evidence and inclusion of diverse information such anecdotal and significant change stories. Bibliometrics can be understood as a range of techniques for assessing quantity, dissemination and content of publications and patents (Guthrie et. al. 2013). Bibliometrics use quantitative analysis to measure patterns of publication and citation with particular focus on journal papers (ibid).

Four themes were identified from the responses on evidence used in evaluation namely:

1. Lessons Learned
2. Bibliometric indicators (number of publications, impact factor, citation indexes,)
3. Economic Analysis
4. Health indicators

Perhaps unsurprisingly bibliometric information was identified as key evidence required for evaluation purposes by agencies and programs that had a specific traditional health research. Some participants noted that bibliometric information individually was not sufficient evidence and needed to be part of package of types of evidence that are required to better understand outcomes and impacts of research interventions. As one participant noted:

Credible evaluation includes some quality of research (peer scholar) and also connection to users of research for example what was helpful (or enlightenment in knowledge use) and grantee perspectives and outcomes.

At the grantee level, there are greater senses that bibliometrics are not grounded or do not adequately capture the complete story of an intervention and its impact. Most grantees and evaluators expressed a greater willingness to explore other forms of evidence particularly significant change stories and anecdotal evidence being cited as particularly useful for understanding lessons learned. As grantee organization take on larger projects and funding and work closer with stakeholders such as governments, bibliometrics are viewed as useful for establishing credibility within global health.

Good Practice: Research Dissemination

A Grantee in South Africa described a project-initiated activity of using drama/theatre to disseminate research results (Home-based AIDS care) in local communities.

At the agency and program level, economic analyses (e.g. cost benefit analysis) were identified as particularly useful for demonstrating the impact of health interventions to board of directors and government funders. The diversity or composition of boards with directors having diverse professional backgrounds meant that in addition to the sought after health indicators, economic analyses were influential in getting directors of boards to understand that type of impact an intervention or research study was likely to have.

While economic analyses were identified as useful for evidence at the regional level, policy influence or demonstrated change in policies was a key type of evidence sought after in evaluations at the regional level.

Evaluation Approach and Frameworks:

Participants found questions about evaluation approaches and frameworks most challenging to answer and often sought clarification on what was meant by approach and framework. Participants identified logical frameworks, outcome mapping tools and change in thinking with regards to results in response to questions on evaluation approaches. The responses suggest variety in understanding of evaluation approaches and no consensus on evaluation frameworks.

The responses to questions on why organizations chose particular frameworks equally elicited diverse responses ranging from donor driven (largely grantees), most applicable to an organization based on literature reviews to comprehensive agency review that pointed to a need for greater accountability and demonstrate results beyond numbers. A participant noted how their particular organization did not conduct *systematic evaluations* and mainly reported on activities. All participants were interested in the impact or return from research interventions and would try and organize their evaluation activities in a way that aims to give an indication of what the impact and return from the research intervention was such that the payback framework appears to be a framework that best captures evaluation approaches at the agency, program and grantee levels.

It is also interesting to note the role of grantees in evaluation activities. At the agency level, it appears grantees have not been included. One participant provided an example of an agency strategy developed for a country (LMIC) that had no involvement of grantee/stakeholders/representatives from the country. Another participant spoke of never having seen or participated in an evaluation of their donor almost ten years into a partnership that has been growing in scope and global health investments. At the program level, participation of grantees is seen as crucial. As a condition of receiving funding, many grantees regularly conduct evaluations and host monitoring teams from various funders. This raises issues of equity and questions on stakeholder involvement such as for who and under what circumstances? The literature is replete with information on the importance of involvement of stakeholders at all levels and throughout the evaluation and project life-cycle. The voices and perspectives of donor recipients are important to informing short and long-term donor strategies and priorities. The omission of these voices and perspectives furthers the perception of evaluations as exercises in financial accountability to donors.

It is interesting to note that participants largely spoke of evaluation of research projects with little in the way of evaluation of programs and agencies given the interest in influencing decision-making and the direction of global health.

The responses from participants were grouped into the following three categories:

1. Evaluation tools
2. Reason for choosing framework
3. Role of grantees in agency level evaluation

Perhaps surprisingly, many respondents were unaware of evaluation practices of other agencies including fellow ESSENCE members. This was particularly telling at the Agency level given the number of forums these Agencies belong to and regularly attend.

One of the deliverables of the present study was to document the nature of evaluation practices, challenges and successes at various levels. A summary of this is highlighted below:

Level	Challenges	Successes
Agency	<ul style="list-style-type: none"> • Measuring impact (scientific and societal) • Prioritization of bibliometrics over other types of evidence • Understanding of how to use research to influence policy and decision making • Use of results and reported activities to influence agency strategic decision-making. • Inter-agency collaboration and cross cutting learning • Limited involvement of grantees in agency evaluation /review • Harmonizing different approaches by agencies and countries in a region to aid with transferability • Weak governance structures at national levels to support EGHR4D particularly for national health systems 	<ul style="list-style-type: none"> • Contribution of M & E to making India Polio Free • Cohort analysis tracking grant recipients 10 years after end of funding • Annual review crucial and push PI to go after goals. Key way to keep track of activities and investments • Impact evaluation demonstrated impacts on health system that senior management needed for endorsement of continued funding
Program	<ul style="list-style-type: none"> • Measuring impact (scientific and societal) • Ensuring relevance of research questions to development/country 	<ul style="list-style-type: none"> • Cross learning through inter-agency evaluation working group (GHRI)

	<p>context</p> <ul style="list-style-type: none"> • Connection between reported activities (Program Managers) and strategic planning (Board) – Ensuring continued funding of activities. <i>Yearly struggle</i> • Limited investment in evaluation capacity 	<ul style="list-style-type: none"> • Operational research and ability to apply lessons in different contexts • Inclusion of grantees in program reviews
<p>Grant/ Project</p>	<ul style="list-style-type: none"> • Measuring impact (scientific and societal) • Donor and stakeholder relationship/engagement. Is grantee able to raise challenging questions? How receptive is donor or govt. rep? - <i>"Negotiating the evaluation"</i> • Ensuring relevance of research questions to development/country context • Limited investment in evaluation capacity for practiners and staff of projects – <i>Evaluation still perceived as an examination of individual performance, still fearful of the evaluation</i> • Need for more south-to-south EGHR4D specific collaboration 	<ul style="list-style-type: none"> • Flexibility and willingness to consider alternative evidence beyond bibliometrics • Participatory approaches involving key stakeholders • Equity agenda • Increasing interest and number of forums to share GHR4D practices and experiences e.g. 2014 Cape Town Conference

Attachment 6b(i) Interview guide for agency representatives

Introduction

Thank you for agreeing to participate in this interview today.

The overall goal of this project is to identify practices and approaches to evaluating GHR4D among different agencies and their grantees, highlighting common successes, challenges and innovations. The purpose of the interview today is to understand how [organization name] evaluates GHR4D, including policies, practices and methods; and successes, challenges and innovations. In this interview we are particularly interested in learning from your experience and insights as a project director/manager about the role that evaluation plays in your organization at a strategic level (E.g., informing decisions and planning about mid- to long-term priorities and investment with regard to GHR4D; demonstrating achievement/progress to key stakeholders; informing the development of organization/GHR4D-wide evaluation policy, procedures and support infrastructure; assessing your organization's contribution to GHR4D). Later in the project we hope to interview staff at individual programme and project levels (and possibly grantees too), to find out about evaluation practices at those levels, and in this interview I will ask you for suggestions about who to interview at these levels.

Verbal Consent

With your permission, I will digitally (audio) record this interview so as to not miss any information. Please be assured that the information you provide will be held in strictest confidence at all times. You may choose to end the interview at any time and you have the right not to answer any question. All interview data will be anonymised in our reports and publications.

Do I have your permission to audio-tape this interview?

YES: _____ No: _____

Do you have any questions before we begin?

General background

- What is your role in [organization name] and how long have you served in this role??
- How long have you been with [organization name]?
- What are the key results your organization seeks to achieve in GHR4D?
 - Does it actively seek to:
 - Advance knowledge?
 - Build research capacity?
 - Inform decision making by managers and policymakers in the health system?
 - Produce health benefits
 - Produce wider socio-economic benefits?
 - Other: _____
- Can you tell me about your organization's GHR4D portfolio?
 - Characteristics /type of programs?
 - Size of programs \$

Evaluation approaches at the strategic level

- What purposes does evaluation serve for your organization at a strategic level (E.g. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure)?
- What sorts of evidence are useful for these purposes??
- How does your organization organize its evaluation activities so as to best serve those purposes in each of the above areas that it is actively involved in (i.e. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure)?
- What specific evaluation frameworks or approaches are used at a strategic level?
 - Why were these chosen/developed?
 - How were they identified/developed?

- Can you provide me with documents describing relevant policies, structures, processes, frameworks/approaches and evidence?

Challenges, learning and good practices + pointers to future interviewees

- What challenges is your organization experiencing with regard to GHR4D evaluation at a strategic level (E.g. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure) ?
 - Please give an example to illustrate these challenges?
- What aspects of GHR4D evaluation at a strategic level (E.g. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure) have been particularly successful?
 - Please give an example to illustrate these successes
- What are your organization's current priorities for evaluation with regard to GHR4D at a strategic level (E.g. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure)?
 - How were these priorities determined?
 - Which stakeholders were involved?
 - Who in the organization would be best placed to tell us about what is being done to deliver on these priorities?
- With regard to the GHR4D programmes* that you mentioned earlier:
 - What approaches are being used to evaluate them?
 - Which stand out because their evaluation has proved either particularly useful or particularly challenging?
 - Can you provide me with any documents about these evaluation activities?
- Who would be a good person for us to contact to find out more about current and past evaluation of these programmes or projects?
- What is your organization's policy with regard to the involvement of grantee organizations in evaluation of GHR4D at a strategic level (E.g. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure)?
 - Does your organization support grantee organizations, countries or regions to use evaluation at a strategic level with regard to their GHR4D?
 - If Yes:
 - How does your organization do this?
 - What successes and challenges have their been?
- How does your organization compare with others with regard to its evaluation of its GHR4D at a strategic level (E.g. informing strategic GHR4D priorities and investments; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure)?
 - Is your organization a member of any benchmarking or learning networks?
 - Which organizations do you think your organization could learn most from?
 - Why?
 - Who would be a good person for us to contact to find out more?

Closing the dialogue

- Are there any other important points that we haven't covered?
- Would you like to receive a copy of our report in due course?

Thank you again for participating in this interview.

* A programme will coordinate, direct and oversee the implementation of a set of related projects and activities (multiple grantees) over a long period of time) in order to deliver outcomes and benefits related to the organization's strategic objectives (e.g. wider socio economic benefits). The realization of outcomes will occur as a result of the programme work of an organization in cooperation with other GHR4D partners. A project (funded research) is usually of shorter duration and with a set of very specific goals and parameters.

Attachment 6b(ii) Interview guide for program level representatives

Introduction

Thank you for agreeing to participate in this interview today.

The overall goal of this project is to identify practices and approaches to evaluating GHR4D among different agencies and their grantees, highlighting common successes, challenges and innovations. The purpose of the interview today is to understand how [organization name] evaluates GHR4D, including policies, practices and methods; and successes, challenges and innovations. In this interview we are particularly interested in learning from your experience and insights as a programme director/manager about the role that evaluation plays in your organization at a programme level (E.g., demonstrating achievement/progress to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D). Later in the project we hope to interview staff at individual programme and project levels (and possibly grantees too), to find out about evaluation practices at those levels, and in this interview I will ask you for suggestions about who to interview at these levels.

Verbal Consent

With your permission, I will digitally (audio) record this interview so as to not miss any information. Please be assured that the information you provide will be held in strictest confidence at all times. You may choose to end the interview at any time and you have the right not to answer any question. All interview data will be anonymised in our reports and publications.

Do I have your permission to audio-tape this interview?

YES: _____ No: _____

Do you have any questions before we begin?

General background

- What is your role in [organization name] and [program name].
- How long have you been with [organization name]? and [program name]?
- What are the key results your program seeks to achieve in GHR4D ?
 - Does it actively seek to:
 - Advance knowledge?
 - Build research capacity?
 - Inform decision making by managers and policymakers?
 - Produce health benefits?
 - Produce wider socio-economic benefits?
 - Other: _____
- Can you tell me about the health portfolio within your program?
 - Characteristics /type of programs/projects?
 - Size of programs/projects\$

Evaluation approaches at the program level

- What purposes does evaluation serve for your program (E.g., demonstrating achievement/progress to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D)?
- What sorts of evidence are useful for these purposes?
- How does your program organize its evaluation activities so as to best serve those purposes in each of the above areas that it is actively involved in (i.e. demonstrating achievement/progress to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D)?
- What specific evaluation frameworks or approaches are used at a programme level?
 - Why were these chosen/developed?
 - How were they identified/developed?
 - Can you provide me with documents describing relevant policies, structures, processes, frameworks/approaches and evidence

Challenges, learning and good practices + pointers to future interviewees

- What challenges is your organization experiencing with regard to GHR4D evaluation at a programme level (i.e. demonstrating achievement/progress to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D)?
 - Please give an example to illustrate these challenges?
- What aspects of GHR4D evaluation at a programme level have been particularly successful?
 - Please give an example to illustrate these successes
- What are your programme's current priorities for evaluation with regard to GHR4D at a programme level (i.e. demonstrating achievement/progress to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D)?
 - How were these priorities determined?
 - Which stakeholders were involved?
 - Who in the organization would be best placed to tell us about what is being done to deliver on these priorities?
- With regard to the GHR4D programmes* that you mentioned earlier:
 - What approaches are being used to evaluate them?
 - Which stand out because their evaluation has proved either particularly useful or particularly challenging?
 - Can you provide me with any documents about these evaluation activities?
- Who would be a good person for us to contact to find out more about current and past evaluation of these programmes or projects?
- What is your programme's policy with regard to the involvement of grantee organizations in evaluation of GHR4D at the programme level (i.e. demonstrating achievement/progress

* A programme will coordinate, direct and oversee the implementation of a set of related projects and activities (multiple grantees) over a long period of time in order to deliver outcomes and benefits related to the organization's strategic objectives (e.g. wider socio economic benefits). The realization of outcomes will occur as a result of the programme work of an organization in cooperation with other GHR4D partners. A project (funded research) is usually of shorter duration and with a set of very specific goals and parameters.

to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D)?

- Does your programme support grantee organizations, countries or regions to use evaluation at the programme level with regard to their GHR4D?
 - If Yes:
 - How does your organization do this?
 - What successes and challenges have their been?
- How does your organization compare with others with regard to its evaluation of its GHR4D at a programme level (i.e. demonstrating achievement/progress to key stakeholders; informing decision-making; informing the development of organization evaluation policy, procedures and support infrastructure; assessing your programmes' contribution to your organizations work in GHR4D)?
 - Is your organization a member of any benchmarking or learning networks?
 - Which organizations do you think your organization could learn most from?
 - Why?
 - Who would be a good person for us to contact to find out more?
-

Closing the dialogue

- Are there any other important points that we haven't covered?
- Would you like to receive a copy of our report in due course?

Thank you again for participating in this interview.

Attachment 6b(iii) Interview guide for researchers (grantees)

Introduction

Thank you for agreeing to participate in this interview today.

The overall goal of this project is to identify practices and approaches to evaluating GHR4D among different agencies and their grantees, highlighting common successes, challenges and innovations. The purpose of the interview today is to understand how [organization name] evaluates GHR4D, including policies, practices and methods; and successes, challenges and innovations. We are particularly interested in learning from your experience about the use of evaluations in your work and with funders. Later in the project we hope to interview funding grantees to find out about evaluation practices at those levels, and in this interview I will ask you for suggestions about who to interview at project and grantee levels.

Verbal Consent

With your permission, I will digitally (audio) record this interview so as to not miss any information. Please be assured that the information you provide will be held in strictest confidence at all times. You may choose to end the interview at any time and you have the right not to answer any question. All interview data will be anonymised in our reports and publications.

Do I have your permission to audio-tape this interview?

YES: _____ No: _____

Do you have any questions before we begin?

General background

- What is your role in [organization name] and how long have you served in this role?
- How long have you been with [organization name] ?
- What are the key results your organization seeks to achieve in GHR4D ?
 - Does it actively seek to:
 - Advance knowledge?
 - Build research capacity?
 - Inform decision making by managers and policymakers?
 - Produce health benefits?
 - Produce wider socio-economic benefits?
 - Other: _____
- Can you tell me about the health research portfolio you lead?
 - Characteristics /type of projects?
 - Size of programs/projects\$

Evaluation approaches at the project level

- What purposes does evaluation serve for your organization at a project level (E.g. informing decision-making; demonstrating progress/achievement to key stakeholders; informing the development of organization-wide/GHR4D policies, procedures and support infrastructure)?
- What sorts of evidence are useful for these purposes?
- What GHR4D projects¹ are you currently leading-participating in?

¹ A programme will coordinate, direct and oversee the implementation of a set of related projects and activities (multiple grantees) over a long period of time) in order to deliver outcomes and benefits related to the organization's strategic objectives (e.g. wider socio economic benefits). The realization of outcomes will occur as a result of the programme work of an organization in cooperation with other GHR4D partners. A project (funded research) is usually of shorter duration and with a set of very specific goals and parameters.

- What approaches are being used to evaluate them? By whom?
 - Why were these chosen/developed?
 - How were they identified/developed?
- Which stand out because their evaluation has proved either particularly useful or particularly challenging? And why?
 - Can you provide me with documents describing relevant policies, structures, processes, frameworks/approaches and evidence?
- Who would be a good person for us to contact to find out more about current and past evaluation of these projects?
- What results or issues arise when your GHR4D work is evaluated?
- In what ways are project evaluations valued and/or not valued by you and your organization?
 - Do you draw on project evaluation to make decisions about modifications?
 - What kinds of evaluations do you value?

Challenges, learning and good practices + pointers to future interviewees

- What challenges is your organization experiencing with regard to GHR4D evaluation at the project level
 - Please give an example to illustrate these challenges?
- What aspects of GHR4D evaluation at the project level have been particularly successful?
 - Please give an example to illustrate these successes
- What are your projects' current priorities for evaluation with regard to its GHR4D strategies?
 - How were these priorities determined?
 - Which stakeholders were involved?
 - Who in your team would be best placed to tell us about what is being done to deliver on these priorities?
- In terms of the funding your organization receives, what kind of evaluation reports and information do you provide funders?
 - What information do funders consider the most and least important for you to provide?
 - As a funding recipient, what information do you consider the most and least important for you to provide?
 - If you receive funding from more than one funder, what are some similarities and differences with respect evaluation requirements by funders?
 - What is working well and what could be improved in this evaluation feedback loop?
- Do you and your organization commission evaluations?
 - What does that look like?
 - How well has this worked?
 - What is your perspective on grantee-led vs. country-led vs. funder-led evaluations?
- How does your project compare with others with regard to its evaluation of its GHR4D project?
 - Is your organization/project a member of any benchmarking or learning networks?
 - Which organizations/projects do you think your project could learn most from?
 - Why?
 - Who would be a good person for us to contact to find out more?

Closing the dialogue

- Are there any other important points that we haven't covered?
- Would you like to receive a copy of our report in due course?

Thank you again for participating in this interview.

Attachment 7 – Additional Resources & Background readings

- Banzai R, Moja L, Pistotti V et al (2011). Conceptual frameworks and empirical approaches used to assess the impact of health research: an overview of reviews. *Health Research Policy and Systems*; **9**: 26
- Boutilier Z, Daibes I, Di Ruggiero E. (2011) Global health research case studies: lessons from partnerships addressing health inequities. *BMC Int Health Hum Rights*. **11** (Suppl 2):S2
- Boyd A, Cole DC Cho D-B et al (2013). Frameworks for evaluating health research capacity strengthening: a qualitative study *Health Research Policy and Systems* **11**:46
- Cheung, A., Lavis, J., Hamandi, A., El-Jardali, F., Sachs, J., Sewankambo, N. & Team, K.-T.P.E. (2011) Climate for evidence-informed health systems: A print media analysis in 44 low- and middle-income countries that host knowledge-translation platforms. *Health Research Policy and Systems*, **9**, 7.
- Cole, D., Boyd, A., Aslanyan, G. & Bates, I. (2014) Indicators for tracking program to strengthen health research capacity in lower- and middle-income countries: a qualitative synthesis. *Health Research Policy and Systems*, **12**, 17
- Crichton, J. & Theobald, S. (2011) Strategies and tensions in communicating research on sexual and reproductive health, HIV and AIDS: a qualitative study of the experiences of researchers and communications staff. *Health Research Policy and Systems*, **9**, S4.
- Earl, S., & Carden, F. (2002). Learning from complexity: the International Development Centre's experience with Outcome Mapping. *Development in Practice*. **12** (3-4): 518-524
- Frenk J, Gómez-Dantés O, Moon S. (2014). From sovereignty to solidarity: a renewed concept of global health for an era of complex interdependence. *Lancet* **383**: 94-97
- Grazier, K.L., Trochim, W.M., Dilts, D.M. & Kirk, R. (2013) Estimating Return on Investment in Translational Research: Methods and Protocols. *Evaluation & the Health Professions*, **36**, 478-491.
- Hamel, N. & Schrecker, T. (2011) Unpacking capacity to utilize research: A tale of the Burkina Faso public health association. *Social Science & Medicine*, **72**, 31-38.
- Hanney, S. & Gonzalez-Block, M. (2013) Organizing health research systems as a key to improving health: the World Health Report 2013 and how to make further progress. *Health Research Policy and Systems*, **11**, 47.

- Koplan JP, Bond TC, Merson MH et al (2009) Toward a common definition of global health. *Lancet* **373**:1993-95
- Kothari, A., Edwards, N., Hamel, N., & Judd, M. (2009). Is research working for you? Validating a tool to examine the capacity of health organizations to use research. *Implementation Science*, **4**(1), 46.
- Landes, M., Lettow, M.V., Cataldo, F., Chan, A., Barr, B., Harries, A. & Bedell, R. (2013) Building a national direction for research in the prevention of mother to child transmission of HIV: results from a national prioritization initiative in Malawi. *Health Research Policy and Systems*, **11**, 40.
- Lange, L. & Luescher, T.M. (2003) A monitoring and evaluation system for South African higher education : conceptual, methodological and practical concerns : perspectives on higher education. *South African Journal of Higher Education*, **17**, p.82-89.
- Lemay MA, Sa C. The use of academic research in public health policy and practice. *Evaluation Research* **23**:79-88
- Masipa, M.D. (2011) *A framework for the evaluation of research in South African Higher Education Institutions: conceptual and methodological issues*. Vol. PhD. Stellenbosch: University of Stellenbosch.
- Morrison V., Keeling M., Jacques M. Wicked Problems and Public Policy. National Collaborating Centre for Healthy Public Policy – Institut national de santé publique du Québec. Montréal (Québec), 2013. Accessed 2014 June 17 at http://www.ncchpp.ca/130/publications.ccnpps?id_article=927
- Orem, J.N., Marchal, B., Mafigiri, D., Ssenogooba, F., Macq, J., Da Silveira, V.C. & Criel, B. (2013) Perspectives on the role of stakeholders in knowledge translation in health policy development in Uganda. *BMC health services research*, **13**, 324.
- Ostlin P, Schrecker T., Sadana R. and others: (2011). Priorities of Research on Equity and Health: Towards an Equity-Focused Health Research Agenda. *PLoS Medicine*, **8** (11): e1001115. Doi:10:1371/journal.pmed.1001115
- Roberts, D. (2013) *Managing Knowledge Results: An Exploration of the Systems and Practices of Development Agencies*. World Bank Institute.
- Smits, P. & Denis, J.-L. (2014) How research funding agencies support science integration into policy and practice: An international overview. *Implementation Science*, **9**, 28.

Sombie, I., Aidam, J., Konate, B., Some, T. & Kambou, S. (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.

Stephen C, Daibes I. Defining features of the practice of global health research: an examination of 14 global health research teams. (2010). *Global Health Action* **3**:5188 - DOI: 3402/gha.v3i10.5188

Terry RF & van der Rijt T. (2010). Overview of research activities associated with the World Health Organization: results of a survey covering 2006/07. *Health Research Policy and Systems* **8**:25

Thornhill, J., Judd, M., & Clements, D. (2009). CHSRF knowledge transfer: (Re) introducing the self-assessment tool that is helping decision-makers assess their organization's capacity to use research. *Healthcare Quarterly*, **12**(1), 22e24.

UKCDS (2013) *Why and what? Motivations for evaluating impact and tools for framing right questions*. London.

Viergever, R., Olifson, S., Ghaffar, A. & Terry, R. (2010) A checklist for health research priority setting: nine common themes of good practice. *Health Research Policy and Systems*, **8**, 36.

World Health Organization (2008). Closing the gap in a generation: health equity through action on the social determinants of health. Commission of Social Determinants of Health Final Report.

Resources related to Approaches and Frameworks:

Barnett, C. & Gregorowski, R. (2013) *Learning about Theories of Change for the Monitoring and Evaluation of Research Uptake*. In: IDS PRACTICE PAPER IN BRIEF. Institute of Development Studies, Brighton.

Brutscher, P.-B., Wooding, S. & Grant, D.J. (2008) *Health research evaluation frameworks: an international comparison*. RAND EUROPE.

Dilts, D.M. (2013) A "Three-Plus-One" Evaluation Model for Clinical Research Management. *Evaluation & the Health Professions*, **36**, 464-477.

European Commission (2011) *EVAL-HEALTH: Developing and testing of new methodologies to monitor and evaluate health-related EU-funded interventions in cooperation partner countries*.

Guinea, J., Sela, E., Gómez, A.J., Jaramillo, H., Gallego, J.M., Patiño, A., Latorre, C., Mangwende, T., Ambali, A., Nyirenda-Jere, T., Seke, L., Ngum, N., Srivanichakorn, S., Thepthien, B. & Putthasri, W. (2013) *Impact Oriented Monitoring (IOM): A new methodology for monitoring and evaluation (M&E) of international public health research projects funded by the 7th framework program of the European Commission*. In: Evaluation of STI policies, instruments and organizations: New horizons and new challenges, Vienna.

Guthrie, S., Wamae, W., Diepeveen, S., Wooding, S. & Grant, J. (2013) *Measuring research: A guide to research evaluation frameworks and tools*. RAND Europe.

Kok, M. & Schuit, A. (2012) Contribution mapping: a method for mapping the contribution of research to enhance its impact. *Health Research Policy and Systems* *C7 - 21*, **10**, 1-16.

TDR WHO (2014) *Implementation Research Toolkit: Workbook*. TDR WHO, Geneva.
Thornton, I. & Shaw, A. (2013) *Evaluating the impact of research programs*. UK Collaborative on Development Sciences (UKCDS), London.

Resources related to the use of evaluation findings:

Hallam, A. & Bonino, F. (2013) *Using evaluation for a change: Insights from humanitarian practitioners*. In: ALNAP Study. ALNAP/ODI, London.

Mayne, J. (2014) Issues in Enhancing Evaluation Use. In: *Enhancing Evaluation Use: Insights from Internal Evaluation Units* (ed. by Laubli-Loud, M. & Mayne, J.). Sage, Thousand Oaks, CA.

Ramalingam, B. (2011) *Learning how to learn: eight lessons for impact evaluations that make a difference*. In: ODI Background Note. ODI, London.

Rudland, E. (2011) *The role of the client in making an evaluation influential: the case of a multi-sectoral evaluation of HIV programs in Papua New Guinea*. In: Australasian Evaluation Society Conference 2011.

Santamaria, M., Schnur, A. & Thapa, D. (2014) Evaluation Policy and Practice in a Changing Environment: Evolution of the Evaluation Function in the World Health Organization (WHO). In: *Enhancing Evaluation Use: Insights from Internal Evaluation Units* (ed. by Laubli-Loud, M. & Mayne, J.). Sage, Thousand Oaks, CA.

Taut, S. (2007) Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation*, **28**, 45-59.

Tennant, S. (2010) Maximizing evaluation influence in an international development donor agency. *Evaluation Journal of Australasia*, **10**, 11-22.