

QUALITATIVE RESEARCH TO ENHANCE THE EVALUATION OF RESULTS-BASED FINANCING PROGRAMMES: THE PROMISE AND THE REALITY

103670

DISCUSSION PAPER

FEBRUARY 2016

Fabian Cataldo
Karina Kielmann



WORLD BANK GROUP
Health, Nutrition & Population

**QUALITATIVE RESEARCH TO ENHANCE THE
EVALUATION OF RESULTS-BASED FINANCING
PROGRAMMES:**

The Promise and the Reality

Fabian Cataldo & Karina Kielmann

February 2016

Health, Nutrition and Population (HNP) Discussion Paper

This series is produced by the Health, Nutrition, and Population Global Practice. The papers in this series aim to provide a vehicle for publishing preliminary results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

For information regarding the HNP Discussion Paper Series, please contact the Editor, Martin Lutalo at mlutalo@worldbank.org or Erika Yanick at eyanick@worldbank.org.

RIGHTS AND PERMISSIONS

The material in this work is subject to copyright. Because The World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Any queries on rights and licenses, including subsidiary rights, should be addressed to the Office of the Publisher, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2422; e-mail: pubrights@worldbank.org.

© 2016 The International Bank for Reconstruction and Development / The World Bank
1818 H Street, NW Washington, DC 20433
All rights reserved.

Health, Nutrition and Population (HNP) Discussion Paper

Qualitative Research to Enhance the Evaluation of Results-Based Financing Programmes: The Promise and the Reality

Fabian Cataldo^a & Karina Kielmann^{a,b}

^a London School of Hygiene and Tropical Medicine, London, UK

^b Institute for Global Health and Development, Queen Margaret University, Edinburgh, UK

Paper prepared for the Health Results Innovation Trust Fund
World Bank, Washington, DC, February, 2016

Abstract:

This Discussion Paper presents the approach, findings, and recommendations from a desk review of the qualitative research conducted within Results-Based Financing programmes (RBF) under the Health Results Innovations Trust Fund (HRITF). The review included 17 studies conducted in Benin, Burundi, Cameroon, DRC, Ethiopia, Haiti, Kenya, Kyrgyzstan, Nigeria, Rwanda, Tajikistan, Tanzania, Zambia, and Zimbabwe. The studies reveal a body of high quality work that is consistent with the conceptual framework of RBF schemes, supported by political will, resources, and research capacity. Strengthening the added value of qualitative inquiry in on-going and future qualitative studies may be enabled by small shifts in thinking and practice, in line with a qualitative research paradigm. First, in order to better ground research in an existing country and system specific context, some interrogation of constructs and posited relationships in the existing conceptual framework for intervention/evaluation may be required. Second, to enable more in-depth and richer data that documents working practices and relations under RBF schemes, training of local researchers should place stronger emphasis on entry to the field, gaining trust, building rapport, and sustaining a dialogue with key informants. Third, smaller, more intensive and focused studies targeting fewer sites and smaller samples - but addressing a wider range of methods and informants within the health system - are likely to yield richer data that can support the understanding of how health workers and managers are responding to schemes, and what impact schemes have on service volumes and outputs.

Keywords: Results-based financing, Qualitative, Methods.

Disclaimer: The findings, interpretations and conclusions expressed in the paper are entirely those of the authors, and do not represent the views of the World Bank, its Executive Directors, or the countries they represent.

Correspondence Details:

Dr. Fabian Cataldo - email: fabiancataldo@gmail.com

Dr. Michael Kent Ranson - email: mranson@worldbank.org

Table of Contents

RIGHTS AND PERMISSIONS	II
FOREWORD.....	V
ACKNOWLEDGMENTS	VI
PART I – APPROACH AND METHODOLOGY:	7
INTRODUCTION.....	7
METHODOLOGY	8
<i>Purpose.....</i>	8
<i>Selection Criteria.....</i>	8
<i>Methods</i>	9
Desk-based review:	9
Semi-structured telephone interviews:	9
<i>Ethical Considerations.....</i>	10
PART II– RESULTS:.....	11
<i>Conceptualisation.....</i>	11
<i>Logistics and Planning</i>	14
Training	14
Sampling	15
Recruitment of Study Participants.....	15
Data Quality Control	16
Ethical Considerations	16
<i>Data Collection.....</i>	16
Study Sites	16
Study Participants.....	17
Methods.....	18
Data Collection Instruments.....	19
Structured questioning.....	20
Open-ended questioning	21
<i>Analysis and Presentation of Results</i>	21
<i>Knowledge Translation.....</i>	23
<i>Limitations.....</i>	24
PART III - CONCLUSIONS	25
PART IV – OPPORTUNITIES AND RECOMMENDATIONS TO STRENGTHEN THE QUALITY OF QUALITATIVE RESEARCH IN RBF STUDIES	26
<i>Conceptualisation of Studies</i>	26
<i>Logistics and Planning</i>	27
<i>Data Collection.....</i>	30
REFERENCES.....	32
APPENDIX: DOCUMENTS INCLUDED IN THE REVIEW	34

FOREWORD

The Health Results Innovations Trust Fund (HRITF), established in 2007 with funding from Norway and the UK, supports the design, implementation, monitoring and evaluation of Results-Based Financing (RBF) programmes with a particular focus on improving maternal and child health outcomes for accelerating progress towards reaching MDGs 1c, 4 and 5. In addition, the HRITF supports activities that build country institutional capacity for RBF and broaden the evidence base for implementing successful RBF mechanisms. A portfolio of rigorous impact evaluations (IEs) have been designed to demonstrate whether these programmes can improve the quantity and quality of health services delivered as well as health outcomes at the population level. As a complement to these IEs, a number of qualitative studies have been carried out to learn about processes of implementation and intermediate components in the causal pathway. Process evaluations and studies of political economy have been supported under the *Learning from Implementation* programme of work. Additionally, many smaller qualitative studies have been conducted in association with HRITF implementation design or impact evaluation. This Discussion Paper focuses on a synthesis review of the qualitative components of RBF-related studies and IEs commissioned by the World Bank under the *Learning from Implementation* programme and through HRITF.

ACKNOWLEDGMENTS

The authors would like to thank the following individuals and institutions for their timely and pertinent contributions to this review: Dr. Clare Chandler, Prof. Janet Seeley, Dr. Dinesh Nair, Dr. Michael Kent Ranson, the Health Results Innovations Trust Fund, the World Bank, the London School of Hygiene and Tropical Medicine, and the study staff and investigators interviewed during this review. The authors are grateful to the World Bank for publishing this report as an HNP Discussion Paper.

PART I – APPROACH AND METHODOLOGY:

INTRODUCTION

Results Based Financing (RBF) is defined as a cash payment or non-monetary transfer made to a national or sub-national government, manager, provider, payer or consumer of health services after predefined results have been attained and verified (World Bank 2013). RBF is an umbrella term that encompasses various types of interventions that target beneficiaries (for example, conditional cash transfers), providers (for example, performance-based financing), and country governments (for example, cash on delivery) (Musgrove 2010).

Project experiences to date show that multiple systems components and their interactions affect health workers' motivation and capacity to implement activities for improved service delivery (World Bank 2014). Likewise, the health-care seeking behaviours of those who are meant to benefit from the health system interventions are determined by many factors, and may be influenced by 'demand-side' incentives.

Programmes utilizing RBF involve complex health systems interventions (Oxman & Fretheim 2008). Bringing in an RBF scheme is inevitably going to impact on issues around financing, governance, and management, including of the health workforce more directly. Dealing with the human impact of these interventions also means dealing with unpredictability. This is where the question of qualitative research comes in: looking at the subjective experiences of health care workers (HCWs) and patients accessing the facilities, the agency of the actors involved, and the way multiple systems components impact on human experience in the context of improved service delivery.

There is consensus that qualitative research methodologies can enhance the understanding of how interventions are implemented within the context of local health systems (Green & Thorogood 2013), and how they do or do not work towards desired outcomes. As a result, the impact evaluations have increasingly used qualitative research methodologies to understand how RBF mechanisms work, and what intermediate components are relevant in the causal pathways between intervention and outcomes.

However, to maximize the potential for qualitative research to generate relevant and meaningful data about processes and mechanisms of effect, research has to be fit-for-purpose, adapted to local context and capacity, asking the right questions, and using appropriate, rigorous, and ethical methods of data collection and analysis. Questions arise in the context of the evaluation framework of RBF projects: What constitutes 'good qualitative research'? Where does it add value to the evaluation of RBF schemes? How is qualitative research best evaluated?

A recent review by Reynolds et al. (2011) identified two dominant narratives in the current literature looking at the quality of qualitative research: one focusing on the outputs and the other on the processes of qualitative research. They recommend that the strengths of both the output-oriented and process-oriented approaches be brought together to create evaluation guidance that reflects core principles of qualitative research, but also responds to expectations of the global health field for explicitly assured quality in research (Reynolds et al. 2011).

In this synthesis review of a few RBF studies supported by HRITF, we do not attempt to assess the validity¹ of results achieved as part of RBF evaluations, as we did not access 'raw' data collected. Instead, we focus our approach on the processes involved in obtaining, managing, and analysing the data.

METHODOLOGY

A desk-based review was conducted focusing on the qualitative components of 17 RBF studies linked to HRITF (see Appendix 1). This review was complemented by a more in-depth investigation of six studies through a case study approach (Stake 1995, Crowe et al. 2011, Yin 2014).

Purpose

The objectives of the synthesis review were to assess the research approach and design, methodological rigor, reporting, and conceptual depth of the qualitative component of studies focusing on RBF; explore opportunities to improve the quality of work conducted as part of the qualitative elements of RBF projects; and offer recommendations to the HRITF team to improve the quality of qualitative projects focusing on RBF initiatives.

Moreover, the additional case study approach aimed to gather further information on the experiences, challenges, and perceived benefits and outputs of conducting qualitative research within six RBF-projects conducted between 2011 and 2015.

Selection Criteria

Projects were selected because they included a qualitative approach or methods focusing on RBF, and were funded directly or indirectly by HRITF.

There were 17 studies included in the synthesis review. These studies were identified in collaboration with the HRITF team. Our review included studies conducted in the following countries: Benin, Burundi, Cameroon, DRC, Ethiopia, Haiti, Kenya, Kyrgyzstan, Nigeria (n=4), Rwanda, Tajikistan, Tanzania, Zambia, and Zimbabwe.

Six of these studies (Cameroon, Ethiopia, DRC, Nigeria, Zambia, Zimbabwe) were selected for a more in-depth assessment. These studies were selected either because they were included in the 'Learning from Implementation' programme (World Bank 2014) or in relation to pragmatic criteria, which included the stage of the research at the time of the review (that is, having reached at least the data analysis stage) and the type of documentation available including, for example, study protocols, Institutional Review Board (IRB) submissions, data collection tools, reports and publications. Note however, that the studies were not directly comparable with respect to the application of qualitative techniques for data collection, management, and analysis. They were developed at different times by different teams, hence do not constitute a coherent portfolio.

¹ Validity here refers to whether the qualitative data adequately and accurately reflect the reality that they were intended to describe.

Methods

In the review of the qualitative components of these projects, a flexible and pragmatic approach was adopted, taking into account the processes involved in carrying out qualitative research as embedded in the broader aims of the projects.

The methods employed were a desk-based review of available documents for each of the 17 studies followed by a semi-structured telephone interview with at least one investigator of six research projects selected for a more in-depth assessment (that is, principal investigator, or co-investigators if the principal investigator was not available).

Review and data collection tools were developed in order to review each study and conduct the interview. The first section of this tool focuses on a descriptive profiling to situate the qualitative elements in relation to the larger research or evaluation. The second section was designed to examine the quality of qualitative research, mainly focusing on study processes prior to, during and after data collection. The third section focused on identifying opportunities to strengthen the place and quality of qualitative research as well as any 'missed opportunities' to strengthen the qualitative component, and identifying where and how the qualitative endeavour may be limited by the overall research approach.

The first and third sections of the tools were developed by the authors of this review. The second section (focusing largely on study processes) was adapted from a tool initially developed for the ACT consortium (Reynolds et al. 2013). Each sub-section was adapted for this review and follows a similar pattern of investigation (i.e. following the study process from conception to dissemination). As part of the tool development, a more in-depth case study approach and interview guide were developed by Karina Kielmann and Fabian Cataldo, and reviewed by all investigators.

Desk-based review:

The desk-based review of documents related to the 17 studies was conducted by one of the investigators and six of these projects (that is, *Learning from Implementation* studies selected for a more in-depth case study approach) were cross-reviewed by a second investigator, discussing any divergent findings.

Semi-structured telephone interviews:

The target group for the telephone interviews included project leads, investigators and/or researchers of six selected projects for a more in-depth case study approach. Following on from the desk-based review, one of the review investigators scheduled a telephone interview with the Principal Investigator, Co-Investigator or project leader for each of the six projects. Five telephone interviews were conducted with project leads and/or investigators, and one investigator was not available for an interview but provided responses in writing.

Each interview lasted between 40 and 60 minutes, and focused mainly on information that was not readily available at the time of the desk-review. Data collected during the semi-structured interviews helped to situate and contextualize specific aspects of the research project, elicit more information in relation to the context of the research, the strengths and weaknesses identified in planning, logistics and implementation of the research, and how the research contributed to understanding processes and outputs for each project.

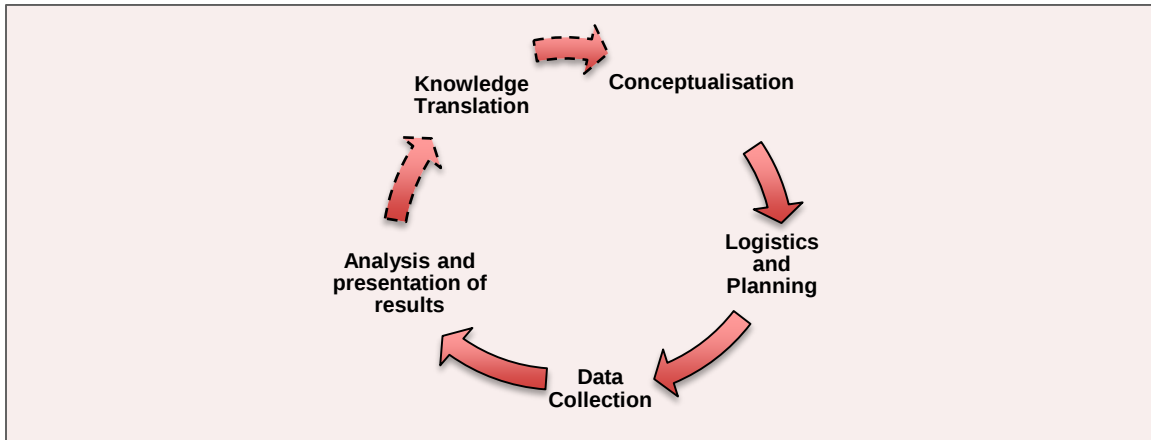
Ethical Considerations

This review was approved by the Ethics Review Committee of the London School of Hygiene and Tropical Medicine (Ref. 8803). Rules for informed consent were respected throughout data collection, and data gained through telephone interviews were anonymized to protect confidentiality of the informants.

PART II– RESULTS:

The results section is organized around the processes that characterize the development and conduct of qualitative research (Figure 1) - namely conceptualisation, planning, data collection, analysis, and knowledge translation (Kielmann et al. 2011). This review focused mainly on issues around conceptualisation, planning and data collection, as limited access was available to the 'raw' data and documentation of the dissemination processes for each study.

Figure 1: Cycle of research enquiry



Source: Authors

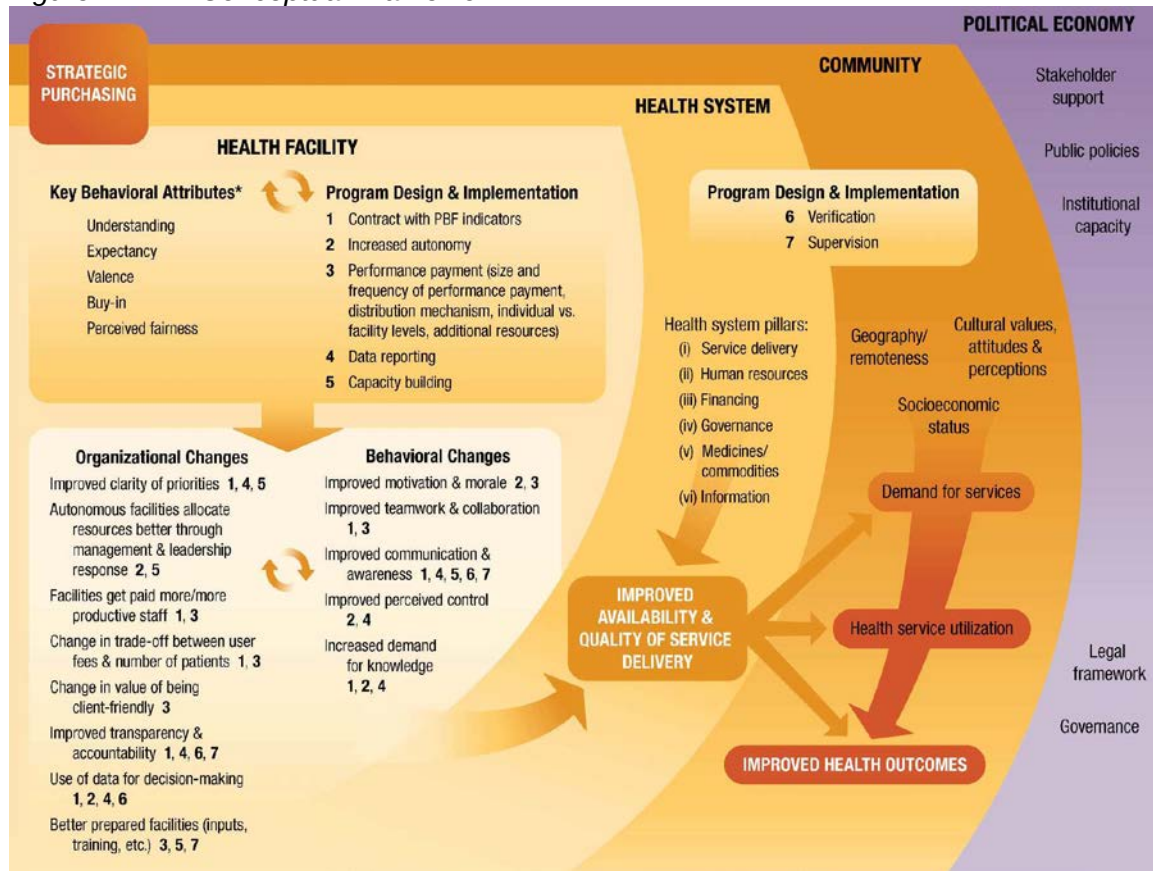
The majority of the RBF studies reviewed were conducted in African countries (n=14); others were located in Tajikistan, Kyrgyzstan, and Haiti. Most were small- to medium-sized studies, generally conducted in the context of an impact evaluation. Overall, studies aimed to explore and document the experiences of health care workers (HCWs), patients, and decision makers in relation to RBF implementation.

In the following sections we cite documents listed under Appendix 1 when referring to a specific example from one of the studies – for example, when referring to the research proposal from the Benin study, we use the citation '1.1: Benin'.

Conceptualisation

All studies were cross-sectional – they tended to use mixed methods, and some were explicitly described as 'case studies' (for example, 9.1: Nigeria, 16.1: Zambia). In several studies (for example, 3.1: Cameroon, 9.1: Nigeria, 17.1: Zimbabwe) the RBF conceptual framework (World Bank 2013; Figure 2) constitutes an explicit starting point for the evaluation and/or qualitative component of the study, although, as stated above, projects did not necessarily adopt a coordinated approach to applying the framework to the research.

Figure 2: RBF Conceptual Framework



World Bank (2013) Using Results-Based Financing to Achieve Maternal & Child Health: Progress Report.

The RBF conceptual framework (World Bank 2013; Figure 2) was recently developed by the HRITF team at the World Bank. It was primarily set up in an effort to rationalise and improve RBF interventions. This model identifies several contextual levels linked to RBF interventions: individual/behavioural, health facility (HF), health system (HS), community and political economy. We used these categories to determine the contextual level at which the 17 studies included in the review were situated. Most of the studies were located at the level of Health Facilities (n=9). Others were focusing on individual behaviours (n=2), community (n=4) and political economy (n=2).²

Several studies included in the review used the RBF conceptual framework to explore the impact of RBF schemes on the health system, and more specifically on HCWs, and in turn how changes in organizational and HCWs' behaviours influence service outputs. The Zimbabwe study (17.1), for example, uses the RBF framework in order to articulate a key study design assumption, which links RBF interventions to health outputs, quality of health services provided, HCWs motivation, and improved access to services in the community.

² Studies focusing on community and political economy were differentiated as follows: by community, we refer to studies that specifically looked at barriers or enablers to uptake of services by beneficiaries. By political economy, we refer to studies that examine the responsiveness and political will amongst key policy donor and high-ranking government officials towards RBF schemes.

In one of the studies conducted in Nigeria, an investigator described how the RBF framework was used and 'customized' during the conceptualisation of the study:

"There are some things I wanted to improve in that conceptual framework: the conceptual framework's focus on how performance changes with PBF [Performance Based Financing]. For example they talk about behavior change at the health centre level but as we are interested in why that behavior change happens, we needed to dig further into details of management practices at the health centre. So instead of focusing on the effects or changes that we can see in the health centre, we wanted to see why those changes actually can happen. So I customized this framework a bit." (9.7: Nigeria)

When adopted in mixed method studies³, qualitative research components can be placed at different stages of a project. Timing is often linked to the purpose of the research. In some of the studies we reviewed, qualitative research was conducted early on as a way of exploring context and understanding on-going processes, for example, in the case of the social assessment conducted in Tanzania prior to RBF interventions (15.1: Tanzania).

In other studies, the research was used primarily to explain quantitative trends observed, for example, variation in the performance of health facilities under RBF schemes, as determined through a set of service output indicators. An example of this is the Nigeria study providing a case study analysis on the best vs. poor performers based on results from an on-going evaluation into HF-related performance linked to RBF pilot schemes (9.1, 9.6: Nigeria):

"We wanted to investigate 'what is going on?' or 'why this is going well, why this is not going well, why there is a difference in performance?'. So qualitative analysis can really unpack these things." (9.7: Nigeria)

Another example from Cameroon illustrates how the qualitative element was intended to generate contextual data in order to complement quantitative data from a larger impact evaluation:

"We felt that doing this qualitative study and making sure that we do cover these different contexts would contribute to the overall impact evaluation and help us once we do the survey to interpret these results." (3.3: Cameroon)

We found no instances in which the qualitative research served to help identify or operationalise the constructs relevant to understanding the impact of RBF schemes on health systems components, and specifically the health workforce. Such research would have helped to understand locally relevant definitions and sources of 'motivation', functional as well as more context and culture-specific dimensions of 'quality' in performance, but also specific constructs to characterise organisational culture, including, for example, dimensions of management and leadership style – hierarchical, vertical,

³By 'mixed methods' studies, we refer to studies which employ a mix of quantitative and qualitative methods to triangulate data collection on a particular topic. As using methods that assess or measure variables of interest through different assumptions, questions, and ways of eliciting data generates different types of data, triangulation allows for a variety of perspectives as well as multiple dimensions of the phenomenon at hand to be explored.

horizontal – that reflect broader societal norms based on gender, occupation, status and so on.

The design of the six *Learning from Implementation* studies appeared to be better supported in terms of resources and more emphasis on qualitative components as part of the overall research approach. This was expressed by several of the investigators interviewed:

“I’m not sure how far we would have actually gone in doing this qualitative study if funding was not made available by HRITF.” (3.3: Cameroon)

In the 17 studies we looked at, only a few were assessing context before implementation of the scheme; the majority were designed to supplement the process or impact evaluation mid-way through implementation, and a few used performance indicators for HFs to examine retrospectively what could explain variations.

Like any other models, the RBF conceptual framework contains assumptions and hypotheses, which frame the methodology. It is instrumental in supporting several of the assumptions made in the qualitative work, for instance the causal relation between the provision of monetary incentives and behaviour change, HCWs motivation, and/or HF overall performance. These assumptions are also reflected in the data collection instruments that were used in the studies reviewed - as discussed in *Section 4.3: Data Collection*.

Logistics and Planning

A critical component of research is logistics and planning leading to data collection itself. Even when the conceptual approach and methodology is sound, it may be impossible to conduct the research as planned because of the specific constraints relating to resources, time, local research capacity, and gaining access and trust in specific sites. We present some considerations in relation to training, sampling, recruitment, data quality and ethics processes.

Training

In terms of capacity to undertake the research, some of the project documents we reviewed included details of training organized as part of the preparation for data collection. Research teams were often structured around task team leaders and senior consultants with technical expertise, but not necessarily in qualitative research. These were supported by one or more local investigators and several field researchers hired on a short-term basis (for example, 4.1: DRC, 3.1: Cameroon, 9.1: Nigeria, 17.1: Zimbabwe). The responsibility for data quality, analysis and reporting often lay with external consultants.

Those collecting qualitative data, that is, research assistants or fieldworkers, were generally trained through short intensive workshops led by one of the international investigators. From the limited documentation available, we noted that the training focused on correct conduct in the field, however this did not necessarily include an emphasis on creating and sustaining relations in the field.

We were pleased to see good practice examples of pilot and pre-testing of instruments, for instance in the Zambia study focusing on HCWs motivation, where detailed training plans and a report were available to review (16.8, 16.9: Zambia).

Sampling

In terms of sampling of research participants, and bearing in mind that the logic of qualitative sampling is different from quantitative sampling, we noted that some of the strategies were not clearly described. While large probability samples attained through random sampling techniques are appropriate in quantitative studies that aim to provide some conclusions regarding how representative a trend is of the wider population, or what differences observed are statistically relevant, qualitative sampling is generally purposive; sample sizes are, on the whole, much smaller, because the aim is to select information-rich cases that can explain variation across a set of informant groups/sets that the researcher deems relevant to the outcome of interest. Other than stratifying the sample of informants according to facility performance (for example 9.3: Nigeria), other features differentiating the sampling of informant groups were rarely mentioned.

An 'ideal' focus group discussion (FGD) has a clear strategy for recruitment of participants based on the research questions, including details of the socio-demographic or professional profile of research participants. Inclusion criteria for FGDs were not always clearly stated in the studies included in the review, and the sampling strategies for community members and gatekeepers were often based on numeric rather than substantive considerations.

An important point is the lack of justification for a relatively large number of respondents interviewed in some of the studies reviewed. A baseline qualitative study in Kyrgyzstan, for example, conducted interviews with 106 individuals (8.1: Kyrgyzstan), while the Cameroon qualitative study had planned to include 168 individual interviews in addition to 67 FGDs (3.1: Cameroon). In some instances, fewer interviews would have been likely to have generated similar findings, and would have been adequate for reporting on the processes. In addition, generating a more manageable volume of data may allow for a more in-depth reflection based on the information gathered.

Recruitment of Study Participants

Protocols for recruitment of participants were in place. A protocol from one of the Nigerian studies, for instance, adequately describes how to recruit research participants (12.3: Nigeria).

Generally, research participants were not directly approached, but fieldworkers went to gatekeepers first or used snowballing approaches to recruit study participants. In several instances, it is not clear how gatekeepers such as 'community leaders' actually represent the communities around them, and how they are defined by the community themselves (for example, 'elected representatives', volunteers, political or religious leaders, etc.).

The inclusion of relatively large samples of study participants has implications for study participant recruitment. Having to recruit large numbers of informants may compromise procedures that are critical for enabling good quality, in-depth data including, for example, adequate processes of rapport building, as well as finding the time and space conducive for the conduct of an in-depth interview and adequate time to do justice to the data.

Data Quality Control

Few studies included detailed information on how to ensure data quality, which involves considerations ranging from creating conditions that are conducive for a good interview, to adequate procedures for recording and reporting of data, as well as reviewing the quality of the data. One study for which we had ample documentation, namely the Zambia project, provided concise, clear protocols on communication (16:5), data collection (16:4), data storage (16:6) and selection of informants (16:7).

Having good protocols in place is not a guarantee, however, for obtaining ‘thick’ data – in reference to Geertz’s “thick description” (1973), that is data that provide rich, in-depth contextual information. This often relies on more intensive interviewing and rapport-building techniques that may be hard to acquire within a relatively short period of time. Ensuring consistently good quality of data may also be contingent on being able to perform ‘reality checks’ on the data periodically to assess whether the data collection guides and instruments as well as the interviewing techniques are eliciting information that is in-depth, coherent, and ‘makes sense’ in the light of what is known about the context.

Hence, another question relevant to this review was when should one assess the quality of the data – often, transcripts are only available for review at the end of data collection rather than mid-way through data collection, allowing for the team to critically review procedures as well as quality of data obtained.

Ethical Considerations

Most studies included in this review have undergone ethical review by the relevant national or international ethics review committees. However, what is mainly being reviewed by institutional review boards are standardized consent forms, data storage protocols, and issues of confidentiality that are concerned with mitigating institutional risk rather than ethical issues faced in the field during data collection.⁴

These ethical issues involve rules of conduct in the field to be adhered to by researchers, but they also relate to the broader framing of projects. Given that researchers were often linked with the evaluation of a specific RBF scheme, it remains unclear from the review of documents and interviews with team members how data collectors and fieldworkers were perceived by informants. An ethical issue which would need to be addressed in the context of these studies is the possible association between RBF schemes and researchers (who may have been perceived as auditors, monitors, or sponsors), which may lead to potential conflict of interest and bias.

Data Collection

Study Sites

The selection of study sites for the qualitative components of the studies was mostly associated with existing RBF schemes and/or impact evaluations. Sites were often selected along the same logic as that used for the impact evaluation, for example, to enable comparisons or uncover reasons for differences in outcomes in relation to location

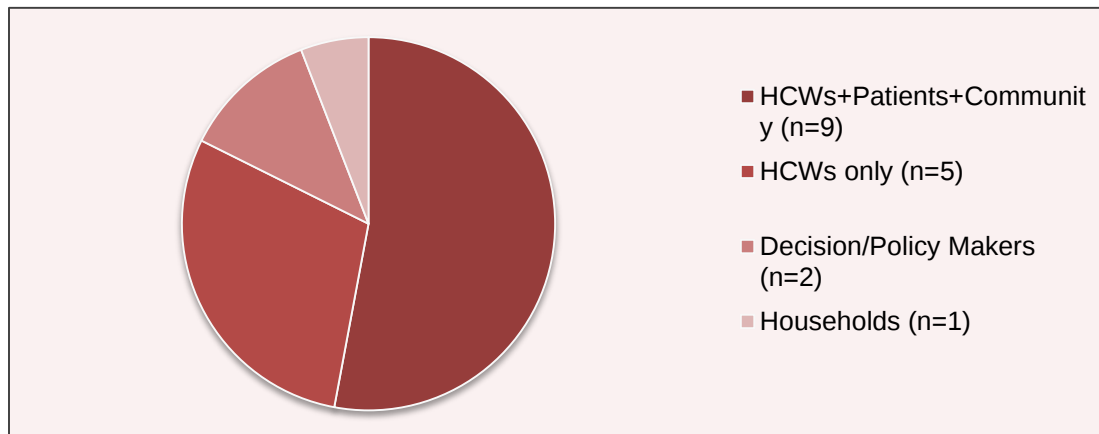
⁴ Guidelines for considerations in building relations with and responsibilities towards research participants can be found in the “Ethical Guidelines for Good Research Practice” published by the *Association of Social Anthropologists of the UK and Commonwealth* (see <http://www.theasa.org/ethics/guidelines.shtml>).

(urban or rural), facility type (dispensary, health centre, or district hospital; for example, 6.1: Haiti) or in relation to performance criteria (high or low performance; for example, 9.1: Nigeria).

Study Participants

Study participants included in the studies reviewed were predominantly frontline health care workers.

Figure 3: Study participants' profile



Source: Authors

In most studies (14 out of 17), HCWs were the main unit of data collection and analysis. In five of these studies, HCWs were the exclusive focus of the qualitative data collection. This is justified through reference to the RBF conceptual framework (Figure 2), which pinpoints HCWs as the locus of behavioural change, mainly in relation to their motivation and performance.

Nine of the 14 studies that focused on HCWs as study participants also included patients and community members. These studies included patients to explore, for instance, what changes could be made to the RBF intervention (for example, 11.1: Nigeria), or factors underlying the success or failure of RBF models on the demand-side (e.g. 13.2: Rwanda). Several studies that included community members did so to elucidate local perspectives and experiences in relation to RBF schemes (for example, 7.3: Kenya), or to evaluate more broadly the potential for RBF in the context of a specific location (for example, 2.1: Burundi).

Relatively few studies included a broader spectrum of health worker cadres that represent or could speak to the dynamics of organizational change. For example, the protocol for one of the Nigerian projects (9.1: Nigeria) included mid-level managers and senior nursing staff as study participants.

Another category of study participant included specific target groups that often served as a pool for key informant interviews (KII). For instance 'community leaders' were study participants in a number of projects. In the Cameroon study (3.2, 3.3: Cameroon), they included the president or leader of the community's women's group and community members who served as the community representative on health centre committees. In

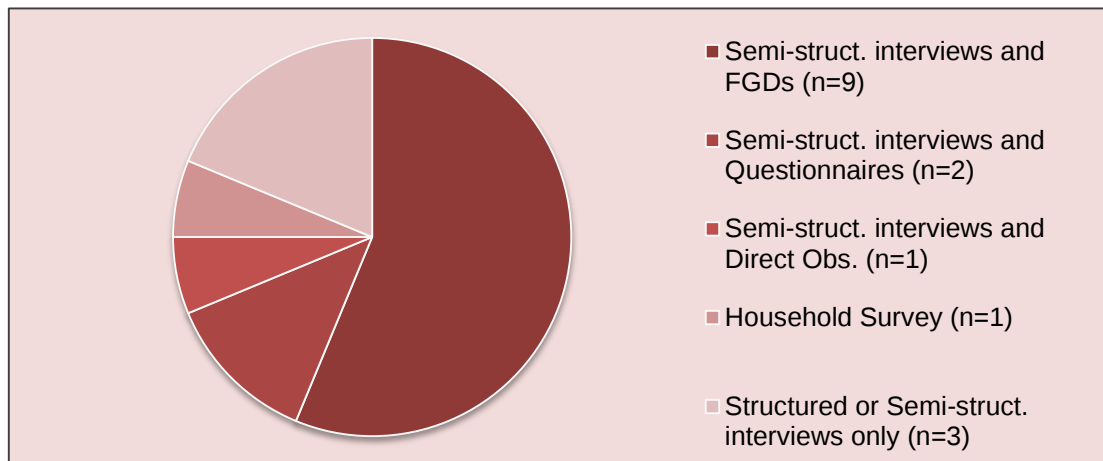
the Rwanda study, they included presidents of community health care workers cooperatives, heads or deputy-heads of HF, and district health officers (13.2: Rwanda). One study (5.2, 5.3: Ethiopia) focused exclusively on high-level stakeholders, who are donor representatives in and outside the country, as well as government officials to assess the pre-implementation context for the RBF scheme.

Methods

A relatively limited range of qualitative methods was used, mainly interview methods relying on reported experience of *what happened* as opposed to more participatory and embedded methods that might document *what is going on*.

The choice of methods seemed to be often made on pragmatic rather than methodological grounds. In most protocols, there was no explicit justification, for example, for the inclusion of FGDs in addition to - or instead of - individual interviews.

Figure 4: Choice of methods



Source: Authors

As illustrated in Figure 4, most studies (16 out of 17) used individual interviews, and these were mainly used in combination with FGDs (n=9). A few studies (n=3) used interviews as a stand-alone method. This was the case for a study relying on stakeholder interviews to examine the pre-implementation context of RBF (5.2: Ethiopia), or KIIs to explore appropriate policy and institutional options for mitigating risks and improving the chances for successful implementation of RBF schemes (10.1: Nigeria). Individual interviews were also used exclusively in one study examining motivation amongst HCWs (16.11: Zambia).

FGDs were often combined with individual interviews, for instance to explore the experience of RBF implementation amongst HCWs and local population groups (3.1: Cameroon). In Haiti, FGDs were used to triangulate data around revenue and human resources from a questionnaire (6.1: Haiti). In Kyrgyzstan, FGDs were used to gather information with women attending primary care facilities to elicit local perception of RBF schemes (8.3: Kyrgyzstan).

In Benin, the research team conducted direct observations (in addition to questionnaires, FGDS, and interviews) to explore the attitude of health providers towards patients during consultations, in addition to data on time and movement mapping within the HFs (1.1, 1.2:

Benin). These observations seemed constrained, however, by the attempt to quantify observed practices, instead of describing patient-providers interactions during patients' visits to the HFs.

Data Collection Instruments

For some, but not all studies, data collection instruments were available for review (see Appendix 1). The guides and instruments reviewed are listed in Table 1.

Table 1: Data collection instruments reviewed

<i>Data collection tool</i>	<i>Number</i>	<i>Studies</i>
Structured questionnaires:	3	1.2: Benin
	1	2.3: Burundi
	1	12.2: Nigeria
	1	17.2: Zimbabwe
Semi-structured interview guides:	1	7.2: Kenya
	2	4.1: DRC
	4	9.2: Nigeria
	2	12.2: Nigeria
	1	13.2: Rwanda
	1	16.2: Zambia
FGD guides:	2	13.2: Rwanda
	1	17.2: Zimbabwe
Observation guide:	1	1.2: Benin
KII guides:	1	10.1: Nigeria
	1	7.2: Kenya
	7	17.2: Zimbabwe

Source: Authors

Most instruments for data collection are semi-structured; the categories of interest are set by researchers in advance, and the instrument contains a mix of closed- and open-ended questions (e.g. 9.2: Nigeria, 13.2: Rwanda, 16.2: Zambia). This is generally also the case for tools intended for in-depth interviews, which ideally should use open-ended questions, or simply a topic guide (list of areas to discuss and probe) to elicit more in-depth information.

The instruments are organised according to a comprehensive set of themes that implicitly reflects the RBF conceptual framework. Hence, for example, questions directed at health workers on motivation relate to the assumed relationship between receiving monetary incentives linked to performance and improved motivation for HCWs (for example, 16.2: Zambia).

Some of the studies reviewed used concepts that derived from the RBF framework, but did not operationalise these further. In principle, it is good practice to use a conceptual framework to guide methodology and research questions, however, it is important to ensure that concepts, especially more abstract ones such as 'autonomy' (for example, 4.2: DRC), 'changes in attrition of HCWs' (for example, 7.2: Kenya), 'performance' (e.g.

3.1: Cameroon), are appropriately operationalized for use in data collection instruments. This means ensuring that the variables chosen to assess the concept are robust and qualitatively or quantitatively measurable, as well as locally meaningful. In one of the studies, there was thoughtful discussion of the literature that could be drawn on to develop indicators for assessing managers' competencies – as good practice, a further question would be how to ensure that the dimensions adapted from available literature/tools would be locally appropriate and applicable (9.1: Nigeria).

Structured questioning

There are substantial issues with survey-based studies that try to understand how people think and behave. For some of the concepts used in the studies, such as 'quality of care', 'job satisfaction', and 'motivation', researchers used Likert scales (Likert 1932) to quasi-quantify informants preferences or ranking of value/importance attached to a particular concept (for example, 12.2: Nigeria, 16.2: Zambia, 17.2: Zambia). While such tools may be powerful when carefully developed and tested for analytical validity in particular settings, under psychometric methodologies, there are also limitations to applying such methods across different settings and without the necessary lead time in developing meaningful measurements. A key limitation is that such scales assume that the concept is understood by the informant in the same ways as the researcher and that a numeric ranking is meaningful. Scales such as these can, in some instances, provide a relatively good sense of how informants evaluate the impact of a particular intervention on their subjective experience – yet results are difficult to compare as concepts such as 'job satisfaction' are not standardised across informants (what makes one person more or less 'satisfied' in relation to his/her job is not the same for the next person). Further, we noted that some instruments used the same scale for a number of questions – which may lead to 'fatigue' and a tendency to gravitate towards the mean on the part of the informant. Additionally, such scales for measuring satisfaction are known to tend towards a positive bias, which limits their utility.

In addition to basing questions around concepts that may or may not be fully operationalised, many of the data collection instruments assumed, rather than probed, relationships between and among concepts, for example, in relation to the effects of specific incentives on motivation and job satisfaction. This is evident in questions such as 'Do you think adding an extra financial incentive will improve your performance?' (16.2: Zambia, 17.2: Zimbabwe); or 'Have you mobilized the community to assist you in increasing the delivery of MCH services' (16.2: Zambia); or 'Avez-vous pu constater des changements dans la structure depuis le début du versement des primes?'⁵ (4.2: DRC).

In these instances, the informants were asked to comment on a presumed situation rather than being given the chance to iterate how they saw the situation and its impact on their working lives. In addition, there were a number of questions that were abstract in nature and did not focus on the concrete experience of the informants, and his or her working practices. Some examples of this were: 'Do you think standards of care at this HF can be improved?' (17.2: Zimbabwe); 'Would you think that the PBF has influenced the experience of patients?' (7.2: Kenya); 'How do you see the motivation of health workers in general?' (12.2: Nigeria).

⁵ 'Have you noticed changes in the facility since incentives started to be paid out?'

Open-ended questioning

The interview guides were structured following a set of topics that covered a number of the areas laid out in the framework. In principle, this is good practice because it ensures consistency between the framework and the data collected. However, in practice, this meant that some of the instruments did not lend themselves to an interview that would encourage informants to speak openly and reflect on what was being said: these instruments lacked a natural or organic flow and tended to move from one topic to the next without adequate links or integration of themes. For instance, the semi-structured interview guide for Zambian HCWs (16.2: Zambia) is divided in clear sections, however these jump from descriptive questions (that is, 'work profile and motivation factor') to perception of challenges (that is, 'understanding challenges and coping up'), and to broad context (i.e. socio-economic and cultural context for MCH care), jumping back to questions about personal satisfaction, motivation and future prospects.

Another point noted was the interspersing of hypothetical questions that asked informants to comment on what they would have liked to see or experience, or how they felt about something that was yet to happen in the future – for instance: 'In five years from now, where do you see yourself?' (17.2 Zimbabwe). Moving between direct questions that ask the informant about a current situation or practice and indirect or rhetorical questions about a future situation may disrupt the flow of the discussion and divert an interview off the main line of questioning. In our experience, informants are generally more at ease, and able to speak freely when asked about their day-to-day concrete working experiences and practices up front. More hypothetical questions about *what should be* are often better placed towards the end of an interview, where they may form a contrasting view or counterpoint to what informants experience on a daily basis.

When asked to reflect on the use of data collection tools, several investigators expressed that more focused tools could have helped to narrow down the analysis:

"To a large extent our set of questions were quite encompassing, we were able to get a lot more than what we had initially thought, [...] one lesson for future studies is to be a bit more streamlined so that we have less... and be less spread across different themes and issues". (17.3: Zimbabwe)

The choice of method for qualitative data collection from these RBF studies may reflect a concern about the volume of data generated by having many research participants. However, we would argue that depth is lost when the data collection tool delimits responses too much.

Analysis and Presentation of Results

We were limited in our review of projects by the documents available for review and in the variation in the degree of completion of each project. In addition, we did not look at primary data but, rather, we looked at how the analysis and results were presented in available study reports. Some of the interviews conducted with study investigators, however, highlight that teams valued the contribution of the qualitative element to contextualise and interpret results:

“By including a qualitative element at least it helps you interpret the results better and really see why the results are what they are, but also, apart from quantitative results and statistical significance, qualitative research can give this very important perspective of how people are feeling and perceiving the intervention on the ground, which is just as important, if not more, than impact at times.” (3.3: Cameroon)

Several studies had strong protocols in place in relation to data management and the steps towards analysis. In many instances, quality assurance processes were in place to supervise transcription and coding:

“During transcription, team members exchanged transcripts for cross-checking and to learn from each other - it was useful that they had a discussion to find out what kind of data they received, this was also used at the analysis level.” (17.3: Zimbabwe)

The investigators of the Cameroon study (3.2: Cameroon), for example, described the data analysis process and how themes and codes were developed with input from the international and national teams (3.3: Cameroon). One of the Nigerian studies also offers a strong rationale for their choice of data analysis approach and describes the process of pattern matching and explanation building (9.3: Nigeria).

There was limited evidence of tools to display data in order to facilitate comparison and contrast across- and within- informant groups. In one preliminary study report (4.4: DRC), matrices were used to summarise data segments across informants. In the Zambia study on HCWs motivation, an analytical framework was generated based on the factors associated with an impact on motivation by respondents (16.11: Zambia).

Although it is difficult to comment on the processes adopted to code and analyse the qualitative data obtained, there is evidence that the data analysis was deductive (researcher-driven), rather than inductive (data-driven) based on the reporting of findings and the ensuing publications, where they were available. This corresponds to the observations made above, namely that the RBF framework was used to guide the research questions, data collection instruments, and the relationships of interest for analysis.

The primary purpose of the qualitative research was to answer specific questions about the likely, intended, or actual impact of RBF schemes on the functioning of health facilities. However, given that the critical focus was on individual HCWs ‘behaviour change’, this meant that other ways of looking at organisational change (for example changes in procedures, roles and responsibilities within teams, processes, and activities intended to boost volume and quality of services) were neglected. Yet, a few studies attempted to look at these dimensions through HCWs coping mechanisms (12.1: Nigeria) and mechanisms leading to improve HCWs performance (1.1: Benin).

In addition to these observations, some of the investigators interviewed during the course of this review also noted that the results from the qualitative elements could have benefited from a more focused approach throughout the research processes, as described in the previous sections. Investigators from the Zimbabwe, Cameroon and DRC studies, for instance, expressed that a limitation in the analysis and presentation of results was the fairly large amount and broad scope of findings:

"We feel that we had specific results, we just feel that we had way too much information. We could have gone on with our analysis, but our report would have been way too big, we could have ended up with two or three different reports." (17.3: Zimbabwe)

"I find the findings very general, this is a nice snapshot of PBF [...], but I'm not sure how much we're learning from it. We might have been a bit ambitious in trying to cover the different impact evaluation groups in different regions and so forth and taking a holistic approach. If we had identified one specific issue and really focused on that, it probably would have been more helpful to resolve that issue. I probably would have narrowed the scope of the research objective." (3.3: Cameroon)

"If more in-depth qualitative research was done during the impact evaluation then the interpretation of the quantitative results would have been more straightforward. When the results came out even the PI said 'we don't know what this means', and that's why we then had to organize a trip there and talk to people, present the results and figure out what happened. If there was a stronger qualitative element in this impact evaluation, the activities we did as an afterthought wouldn't have to be done." (4.7: DRC)

While in some cases, the perceptions regarding the 'bulk' or lack of focused data may be linked to the problem of non-specific instruments, or instruments that were poorly understood or applied, they may also reflect some limitations in the capacity to manage and analyse qualitative data.

Knowledge Translation

This is an area on which we can provide limited discussion because some of the projects reviewed are at different stages of completion, and because documentation related to dissemination plans, for example, were only briefly summarised in study protocols and reports.

All study teams interviewed made plans for local dissemination through workshop meetings, and had already conducted presentations to international audiences. Several studies, as part of their results sections and executive summaries, highlight key recommendations for policy and RBF programme implementation (for example, 3.2: Cameroon, 4.3: DRC, 5.2: Ethiopia, 16.11: Zambia, 17. 1: Zimbabwe). In addition, a few of the studies which were included in the review have already been published in peer-reviewed journals (4.6: DRC, 12.5: Nigeria), or through policy notes/case studies (4.3: DRC, 5.2: Ethiopia).

Investigators highlighted the potential for positive impact from results generated by the qualitative studies to improve interventions and provide locally relevant recommendations:

"This [the study] was a good opportunity to generate evidence specific to the Zimbabwean context that would enable the government to further customize the design of the programme to the local context". (17.3: Zimbabwe)

"[The qualitative study] was a very useful thing to do because we understand more in-depth about what is important or not. Findings are very interesting and actually

can influence what we are doing at this moment in the project and also how we design the project going forward. For example, we clearly understood how the communities really support the health centre [...] that actually didn't come up at the initial hypothesis, so that's something very new and an important finding for the other programme.” (9.7: Nigeria)

Study investigators also described the involvement of local stakeholders during the course of study development, which in relation to knowledge translation may help study results to inform more efficiently both on-going and new RBF schemes:

“There was an engagement from the onset in the planning and during the roll out of the study, which made it very clear for those who are in the leadership locally to understand the nature of the data and the quality of that data.” (17.3: Zimbabwe)

Based on results from the RBF evaluation that took place in DRC, a policy note was developed and its authors attest to its capacity to engage more specifically with local and international policy makers:

“People worldwide have a different understanding as to what RBF is about and how it can be used; some people have a narrow view of RBF that this is only about financing. They tend to omit all the tools and the managerial capacity it can bring to a system, to a health centre [...]. A policy note alone is not enough because it is a lot of hand holding at the country level for the various stakeholders to understand what we mean by this approach and for the different partners to also understand it and see how they can potentially use some of those tools to really achieve some of the intended results [...]. A policy note alone is not sufficient but clearly, in terms of engagement, and captivating people's attention to be willing to sit and talk to you, I think the policy note was instrumental.” (4.7: DRC)

Limitations

A key limitation in our approach lies in the type and availability of documents that were reviewed; in some cases we needed to rely on a few documents to review a project (see Appendix 1). These documents were provided to the review team while the studies were at various stages of completion, which limited our ability to compare some of the outputs, for instance in relation to the presentation of results and dissemination activities.

Secondly, phone interviews were conducted with only one senior member within each study team, and this person was engaged to varying degrees in the fieldwork itself. An important perspective from fieldworkers and data collectors is missing from this approach.

Thirdly, it was difficult to assess issues related to reflexivity, that is the researcher's awareness of how his/her position and status in the field influences the study questions and the data collected, within a single short interview and within the process of reviewing documents. Our approach is therefore limited in relation to information leading to critical reflection on the assumptions, conceptualisation and conduct of research.

PART III - CONCLUSIONS

The inclusion of qualitative research to add depth and value to RBF research and evaluation is timely and highly relevant. There appears to be a will, acceptability and funding as well as growing capacity on the ground to conduct this type of research.

Our review of the qualitative research conducted around RBF programmes to date shows a body of high quality work demonstrating consistency, coherence and rigorous procedures for collecting data that is relevant to answering many questions in the countries where RBF schemes are being implemented.

Through the review of documents provided, we suggest that there are, however, opportunities for maximizing the potential and promise of qualitative research in the context of RBF studies. The body of research reviewed demonstrates that a shift in research paradigm may be required to maximise the potential of qualitative research to inform the operations of the RBF schemes. It is advisable to gain insights, guidance, and on-going support from a social scientist with strong qualitative methodology skills that extend beyond data collection, and include other aspects of qualitative research, for example, social science theory as applied to health –related research, experience in conducting in-depth fieldwork, and capacity for analysis of qualitative data. Where feasible, this should ideally be enabled through partnerships with local research institutions, and more long-term commitment of these partners to the on-going monitoring and evaluation of the RBF schemes.

We further suggest that a more open and flexible approach is adopted vis-à-vis the existing conceptual framework for intervention/evaluation so that some interrogation of constructs and posited relationships can refine the framework to better reflect local health systems and programmatic conditions. In addition, a greater focus on the importance of trust, rapport, and exchange with informants may require the need for a different type of and focus in training and field preparation. Lastly, we suggest that smaller, but more intensive and focused studies are likely to yield richer qualitative data both at the outset of the study as well as in the documentation of processes and mechanism of effect.

PART IV – OPPORTUNITIES AND RECOMMENDATIONS TO STRENGTHEN THE QUALITY OF QUALITATIVE RESEARCH IN RBF STUDIES

In this section, we discuss opportunities to strengthen the quality of qualitative research related to RBF interventions. The implications of key critical points as highlighted for three main stages of qualitative research are discussed, namely first, *conceptualisation* of the research; second, *logistics and planning* and third, *methodology*. We offer brief recommendations for investigational or research strategies to address critical points highlighted throughout our review. We note, however, that these cannot be generic, but must be adapted to the local context, in other words, the specific capacity, skills, study setting, and overall aim and purpose of the qualitative research, as well as in relation to the broader objectives of the evaluation.

Conceptualisation of Studies

In most studies reviewed, the RBF conceptual framework is used to explore the impact of schemes on the health system, specifically on HCWs, and in turn how changes in organizational and HCWs behaviours influence service outputs. Like any other model, the RBF conceptual framework contains assumptions and hypotheses, which frame the methodology in a particular way.

First, this relates to the **context** deemed important in the research. We found that while the broader context of resource allocation within the health sector of respective countries was well-described in relation to health systems capacity to adopt and implement RBF schemes, the more narrow ‘meso-level’ context of management structure and organisational culture within health facilities was assumed, but not explored, although this context is critical for human resources for health (HRH) issues. It is critical to capture, for example, how infrastructural and resource constraints, as well as specific organisational features of health facilities impact on health workers’ motivations and ability to ‘*do the right thing*’, that is, what they were trained or incentivised to do.

Second, the framework places emphasis on specific **actors**. Hence, while recognising that HCWs are embedded within organisations, they are seen as the locus of behaviour change, resulting in a misperception of their capacity for decision-making and agency in isolation from other actors.

Third, the framework delineates a number of **constructs and their relationship** relating to behaviour and its impact on desired outputs. While these are well-specified and allow for good consistency throughout the methodology for many of the studies reviewed, we found that too many assumptions were both implicit and static, limiting the richness and context-specificity of the data that could be obtained in the respective settings.

We suggest that thinking about RBF as a complex health systems intervention, as some recent researchers have (Macq 2011, Witter et al. 2013) will allow for a more dynamic understanding of how RBF schemes impact on health facilities and the staff that man them. This is critical in the field of HRH where concepts of organisational culture, management, and leadership are highly influenced by, and contingent on existing social hierarchies and relations of power, trust, and accountability that are mapped on to the health system.

Witter et al. (2013) provide a framework for monitoring and evaluating the health systems effects of RBF schemes, which suggests that areas for inquiry under processes relating to effects of RBF on service delivery, human resources, governance, and financing are not only complex and dynamic in nature but inter-linked (see also Ssengooba et al. 2012). Understanding changes in 'accountability' or 'autonomy' of facilities frequently involves an in-depth examination of the impact of the initiative on existing relations within the facility – across and between health care worker cadres, as well as between health care staff and users. This suggests a need to consider not only individuals (for example, health workers or managers) as the units of analysis in qualitative studies, but to shift the focus to the social dynamics of 'working relations' between individuals and amongst teams. Closer attention might, for example, be paid to the balance of power and authority within health facilities, communication channels and decision-making hierarchies, as well as sites of cooperation and conflict within teams.

Recommendations 1: Conceptualisation of studies

The following steps can strengthen the conceptualisation of RBF studies:

- A thorough literature review to understand how key concepts and their relationship have been defined and applied in the regional setting, where applicable. While there may not be specific empirical research in the country where the study is being undertaken, it will be useful to look at how dimensions of a concept have been applied and assessed in different settings.
- A period of formative research involving site visits, familiarisation local health institutions, and largely informal interactions with key individuals can allow for a mapping of relevant structural and organisational features that may impact on health workers' behaviour and working relations in respective sites. The site visits serve two critical aims, firstly to facilitate a series of logistic steps - making contact, gaining access, deciding on sampling and recruitment strategies – and secondly, to enable refinement of a context-specific conceptual framework to inform the research.
- The literature review and formative research allow for:
 - a) More focused research questions that resonate with the experiences of health workers and other informants being interviewed and;
 - b) Operationalisation of abstract concepts (especially around organisational and management behaviour) that will help to align tool development more closely with the local system-specific working structures and practices.

Logistics and Planning

In many of the studies reviewed, we saw evidence of good practice in planning research components, through examples of ethical review proposals, information sheets for participants, and in some cases protocols for hiring and training researchers, field conduct, data storage and management. In some cases, standardised protocols for fieldwork procedures are appropriate, as there may be agreed areas that are less dependent on the local context of research.

Having standardised protocols for example, for recruitment of field researchers, division of roles and responsibilities within teams, data recording, management, and quality control, as well as other reporting requirements can be agreed across RBF Team Task Leaders responsible for commissioning or oversight of qualitative research, and permit a certain degree of uniformity across sites. However, we found that there was little documentation of the ethical, social, and practical dimensions of 'real-life' field encounters that have a major bearing on the quality of qualitative research and data collected.

It is important to note the complexity of the field and ensuing demands on fieldworkers. These are not easily transmitted through short training inputs that, by necessity, focus on developing qualitative research techniques and adherence to procedures for data collection, recording, and management. Even if fieldworkers have understood and applied procedures and tools competently in the space of a short training period and pilot study, there are numerous features of conducting qualitative research in low-income health settings that should not be underestimated. These include, for example:

- **Gaining physical access to sites** that may be spread far apart, and in settings where transport, weather conditions, and distance impinge upon set targets of what can be achieved in a day of fieldwork.
- **Being accepted by informants**, establishing contact and gaining respect as a researcher. This involves not only being taken seriously as someone who is legitimate in terms of requesting information (note that features of the researchers' identity such as gender, professional status, and ethnicity may all impact on this interaction) but also being able to competently explain what the purpose of the research is, without compromising the type of data that can be obtained because informants feel that their responses have a direct influence on the continuity/sustainability of the scheme in place.
- **Finding appropriate times and spaces for conducting research** in often crowded, and busy health facilities. The impact of time and space on confidentiality, rapport, and the type of information that can be discussed is considerable.
- **Gaining trust** over repeated encounters is critical for research that will take place over a period of time. It is important to acknowledge and respond to the challenge of informants using a range of means to resist or indirectly refuse interviews (e.g. not turning up for interviews, making fieldworkers wait for long periods of time, delegating the 'interview' to another staff member) because of mistrust, misunderstanding, or simply a perceived waste of time on their part.

These and other dimensions of the research process are likely to have been important for many or indeed all of the projects reviewed, however they were not documented.

We also noted that site selection, sampling, and recruitment of participants often reflect quantitative logic or pragmatism, rather than a methodological rationale. On the one hand, the attempt to cover a wide selection of sites, and achieve large samples for different participant groups reflects perhaps an unspoken assumption that 'more is better' in terms of validity of representativeness and validity of the data obtained. On the other hand, site selection and sampling of particular informant groups appeared to reflect pragmatic choices based in considerations of available time, convenience, and relative facility in access to particular categories of informants.

We suggest that the often vast amount of work related to sampling, recruitment, data collection, and data management and analysis that is evident in many of the studies reviewed is likely to be disproportionate to gains in relevant data. If justified through methodological considerations that take account of locally relevant variation in units of analysis (e.g. context, systems, facilities, health worker cadres), the numbers of sites and sample sizes are likely to be smaller and more meaningful in relation to the research objectives.

Recommendations 2: Logistics and Planning

- Training and developing protocols for fieldwork and data collection with qualitative teams requires more time and attention to a number of specific areas:
 - a) Gaining access to sites, and building respect and trust with key individuals within sites; it is relevant here to create relationships with individuals who may be consulted at later stages of the project (for example, through iterative key informant interviews), who provide feedback on the research process, and who also support local ownership and knowledge translation of the findings.
 - b) Building rapport within teams should not be underestimated, especially when there are a number of research assistants who differ in personality, training, background and other traits that can be a source of comparison, if not friction. It is important, as a field coordinator, to motivate, and bring out the best in research assistants through training and attention to specific strengths and limitations represented on the team.
 - c) Enough time should be spent on emphasising how crucial it is to establish good relations with informants in the field. This goes beyond interviewing techniques and use of the data collection instruments, and extends to social and cultural norms around courtesy, dialogue, empathy, and trust, for example.
- '*Less is more*' – this principle holds for both the selection of sites, as well as for sampling strategies:
 - a) Site selection should reflect adequate qualitative rather than quantitative rationale. In other words, sites for data collection may relate to the need to capture a range of experiences, including those of 'positive' and 'negative' deviants – rather than representativeness of a country or a region.
 - b) Sampling strategies similarly need to reflect a meaningful methodological rationale, in other words, samples for qualitative data collection are often chosen because they relate to a choice of informant type to answer specific kinds of questions. There are also loose guidelines around the recommended sample sizes for specific methods that are defined on the basis of what the method is trying to achieve. For example, a study relying on 'key informant interviews' would consult no more than between 10 to 15 key informants on the average since it is assumed that the number of 'experts' able to provide overview and critical reflection on a topic is finite.

Data Collection

As with site selection and sampling strategies, we observed that the choice of methods in many of the studies was based on pragmatic rather than methodological considerations. The types of methods used were limited in the majority of studies to individual or group interviews. In line with the conceptual framework, data collection is strongly biased towards *reported* perceptions and experiences of frontline HCWs.

These methods may fail to capture *what is happening on the ground* from the different perspectives of the range of actors that are involved in running a health facility and held accountable for service outputs. In instances where other stakeholders (e.g. policy makers, 'community members') were considered, questions arose around the choice, relative contribution, and added value of these informants in terms of their ability to speak to the topic at hand.

We noted that data collection instruments are on the whole, comprehensive and well-structured in terms of the dimensions of the framework they attempt to capture. However, because they are essentially researcher-driven, there is limited opportunity within the instruments for opening up a dialogue or probing further with informants. Despite the fact that most of the data collection tools were deemed 'semi-structured' or 'in-depth', there was limited space to explore topics that are not included in the framework. There are questions around whether informants understood the concepts and assumptions underpinning the interview guides in the same way as the researcher, for example, the perceived relationship between workload, incentives, and performance. Interview guides, while well structured, were not always informant-friendly; with a few exceptions, there were limited instructions or probes included for data collectors/field workers to make the transition between topics, to summarise and check what was being said, or to allow the informant to speak more freely about concrete and personal experience.

Recommendations 3: Data Collection

- Limited but critical use of ethnographic and 'real-time' methods can help shed light on *what is going on*, and can complement reported accounts. These can include, for example, some participant-observation in settings where this is feasible and ethical (e.g. attendance at staff meetings; spending time in waiting areas or at clinic reception areas), using more in-depth interviewing techniques to solicit narrative accounts from both health workers and patients (e.g. '*a typical day...*'; '*describe the last time you managed a patient with x...*', '*tell me about the time you came to the clinic for y...*'), and use of tracking or time-motion studies to better understand pathways and processes of communication, decision-making, and referrals within the facility.
- More focus and depth in data collection instruments that allow informants to 'tell the story' of RBF in more details and in relation to specific themes, using more personal and experiential terms (e.g. eliciting views on working practices, peer relations within health facilities, experience of patient care trajectories).
- Qualitative topic guides may be organised to facilitate a more organic flow of the interview, and involve less direction on the part of the researcher, leaving concepts and assumptions open to inquiry, and the field worker to probe more in-depth along the lines of what is being said, rather than what is 'in print'.
- Standardised and on-going training for data collection teams and field supervisors to ensure consistency in the quality and depth of the data collected. Spending more time to 'pilot' tools whilst being supervised and having access to training resources is essential to ensure that data collection is conducted appropriately for more specialized qualitative techniques, for example in-depth interviewing, probing, and observations.

Due to the limited documents that pertain to data analysis and knowledge transfer that were available for review, we focused our comments regarding methodology on the preparation and use of methods for data collection. However, we offer here additional recommendations in relation to qualitative data analysis and knowledge translation that may be applied to future RBF studies:

Recommendations 4: Analysis and Knowledge Translation

- Data quality assurance requires resources and time. It is important to check not only that data is completely and accurately recorded, but that it is providing rich, context-specific information that helps to inform and add to the understanding of how RBF plays out in different contexts. This needs to be monitored at regular intervals, especially if studies involve larger samples.
- Analysis processes should involve small teams, and allow for feedback from individuals involved in the actual data collection, and for discussion of new, unexpected, surprising or even conflictual accounts. While this 'collaborative' process of coming up with a sound thematic framework that is the basis of a coding system may take more time, it will help to provide direction and focus in the organisation and interpretation of textual data which can amount to hundreds of transcribed pages.
- Consultation and engagement with local stakeholders from the onset of RBF studies can help refine and focus the study results, and may promote the uptake of study results more efficiently at country level.

REFERENCES

- Crowe S., Cresswell K, Robertson A, Huby G, Avery A, Sheikh A. 2011. *The Case Study Approach*. *BMC Med Res Methodol* 2011, 11:100.
- Daly, J., Willis K, Small R, Green J, Welch N, Kealy M, and Hughes, E. 2007. *A Hierarchy of Evidence for Assessing Qualitative Health Research*. *Journal of clinical epidemiology*, 60(1), 43-49.
- Dixon-Woods M., Shaw R. L, Agarwal S, and Smith J. A. 2004. *The Problem of Appraising Qualitative Research*. *Quality and Safety in Health Care*, 13(3), 223-225.
- Eichler R., Levine R. 2000 *Performance Incentives for Global Health: Potential and Pitfalls*. Washington: Center for Global Development, Performance-Based Incentives Working Group; 2009.
- Geertz, C. 1973. Thick Description; Toward an Interpretive Theory of Culture. *The interpretation of cultures*, 3-30.
- Green, J., and Thorogood, N. 2013. *Qualitative Methods for Health Research*. Sage.
- Kielmann, K., Cataldo F., and Seeley J. 2011. Introduction to Qualitative Research Methodology. *Department for International Development (DFID)*.
- Likert, R. 1932. A Technique for the Measurement of Attitudes. *Archives of Psychology*.
- Macq, J. and Chiem, J-C. 2011. *Looking at the Effects of Performance-based Financing through a Complex Adaptive Systems Lens*. *Bull World Health Organ* 89(9): 699–700. doi: [10.2471/BLT.11.089920](https://doi.org/10.2471/BLT.11.089920)
- Musgrove, P. 2010. Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and a Short Glossary. Results-Based Financing for Health, Center for Global Development, World Bank, Washington, DC.
- Oxman A. D., and Fretheim A. 2008. An Overview of Research on the Effects of Results-based Financing.
- Reynolds J., Kizito J, Ezumah N, Mangesho P, Allen E., and Chandler C. 2011. *Quality Assurance of Qualitative Research: A Review of the Discourse*. *Health Research Policy and Systems*, 9(1), 43.
- Reynolds J., Naiga S, Taaka L, and Chandler C. 2013 *Quality Assessment and Strengthening of Qualitative Research: An Example Protocol*. Act Consortium. <https://globalhealthtrials.tghn.org> - accessed February 2015:
- Stake R. 1995. *The Art of Case Study Research*. London: Sage Publications.
- Ssengooba F, McPake B, Palmer N. 2012. *Why Performance-based Contracting Failed in Uganda – An ‘Open-box’ Evaluation of a Complex Health System Intervention*. *Soc Sci Med* 2012; 75: 377-83. doi: 10.1016/j.socscimed.2012.02.050

Witter, S., Toonen J, Meesen B. et al. 2013. *Performance-based Financing as a Health System Reform: Mapping the key Dimensions for Monitoring and Evaluation*. *BMC Health Services Research* **13**:367. doi:10.1186/1472-6963-13-367

Witter S., Fretheim A, Kessy F. L, Lindahl A. K. 2012. *Paying for Performance to Improve the Delivery of Health Interventions in Low- and Middle-income Countries*. *Cochrane Database Syst Rev* 2012; 2. doi: 10.1002/14651858.CD007899.pub2

World Bank. 2013. *Using Results-Based Financing to Achieve Maternal & Child Health: Progress Report*, World Bank, Washington, DC.

_____. 2014. *Resource Based Financing; A Smarter Approach to Delivering More and Better Reproductive, Maternal, Newborn, and Child Health Services*. Health Results Innovation Trust Fund, World Bank, Washington, DC.

Yin R. K. 2014. *Case Study Research: Design and Methods*. London: Sage publications.

APPENDIX: DOCUMENTS INCLUDED IN THE REVIEW

Country	Study	Documents Reviewed
Benin	<i>Evaluating the impact of PBF on health workers' performance in Benin: a mixed method study</i>	1.1 Study proposal (English) 1.2 Study protocol (French) 1.3 Information sheet and Consent forms (French) 1.4 Data collection tools (French)
Burundi	<i>Etude d'impact du FBP Nutrition au niveau de la communauté au Burundi</i>	2.1 Protocol (French) 2.2 Information sheet and consent forms (French) 2.3 Questionnaires for HCWs (French)
Cameroon	<i>Cameroon PBF Impact Evaluation</i>	3.1 Qualitative study proposal (English) 3.2 Qualitative midline study report (English) 3.3 Interview with one of the PIs by phone (English)
DRC	<i>Performance-Based Financing for Health in Haut-Katanga</i>	4.1 Interview tools (French) 4.2 Consent forms (French) 4.3 Policy note (English) 4.4 Field notes (French) 4.5 Study report (English) 4.6 Peer reviewed publication (English) 4.7 Interview with policy note authors by phone (English)
Ethiopia	<i>The World Bank's Program for Results Instrument: A Case Study of its Preparation in Ethiopia</i>	5.1 Terms of Reference (English) 5.2 Case study report (English) 5.3 Email correspondence with the PI (English)
Haiti/Tanzania	<i>Etudes sur les Ressources Humaines et les Dépenses de Santé Dans les Départements du Nord-Est, Plateau Central et Nord-Ouest, Haiti/ Results from a Rapid Assessment of Health Expenditure and Health Worker Motivation, Satisfaction, and Compensation in Shinyanga Region in Tanzania</i>	6.1 Study report (French) 6.2 Results summary (French) 6.3 Study report from rapid assessment (English)
Kenya	<i>Evaluation of performance-based finance (PBF) pilot in Samburu County, Kenya</i>	7.1 Study protocol (English) 7.2 Interview guides (English) 7.3 Study report (English) 7.4 Terms of reference (English)
Kyrgyzstan	<i>Maternal & Neonatal Health Care in the Kyrgyz Republic Baseline Qualitative Study</i>	8.1 Preliminary results report (English) 8.2 Terms of reference (English) 8.3 Study report (English)
Nigeria	<i>Nigeria State Health Investment Project (NSHIP) Health Facility Performance Analysis</i>	9.1 Study proposal (English) 9.2 Interview guide (English) 9.3 Study report (English) 9.4 Observation notes (English) 9.5 Consent forms (English) 9.6 Preliminary findings summary (English) 9.7 Interview with one of the PIs by phone (English)
Nigeria	<i>Political Economy and Institutional Assessment for Results Based Financing for Health</i>	10.1 Study report (English)
Nigeria	<i>Formative Research on Demand-Side Interventions Under Nigeria State Health Investment Project (NSHIP)</i>	11.1 Study report (English)

Nigeria	<i>Assessing health workers' revenues and coping strategies in Nigeria — a mixed-methods study</i>	12.1 Terms of reference (English) 12.2 Study proposal (English) 12.3 Fieldwork manual (English) 12.4 Study report (English) 12.5 Peer reviewed publication (English)
Rwanda	<i>Impact Evaluation of Rwanda Community Health Performance-based Financing (Q-CPBF)</i> <i>"A qualitative study to understand findings from previous and ongoing quantitative survey"</i>	13.1 Terms of reference (English) 13.2 Study proposal and interview guides (English)
Tajikistan	<i>Impact Evaluation of the Tajikistan Results Based Financing Project</i>	14.1 Terms of reference (English)
Tanzania	<i>Social Assessment Report- Tanzania Results-based Financing (RBF) Project</i>	15.1 Study report (English)
Zambia	<i>Determinants of motivation among rural health workers in Zambia and the perceived quality of maternal and child health care</i>	16.1 Terms of reference (English) 16.2 Interview guides (English) 16.3 Consent forms (English) 16.4 Interview protocols (English) 16.5 Communication protocols (English) 16.6 Data storage protocols (English) 16.7 Sampling protocols (English) 16.8 Training plans (English) 16.9 Training report (English) 16.10 Ethics proposal submission (English) 16.11 Study report (English) 16.12 Interview with one of the PIs by phone (English)
Zimbabwe	<i>Process Monitoring and Evaluation of Zimbabwe's Results Based Financing (RBF) Project: The Case of Mazowe, Chipinge, Zvishavane, Binga and Kariba Districts</i>	17.1 Study report (English) 17.2 Interview guides (English) 17.3 Interview with one of the PIs by phone (English)

This Discussion Paper presents the approach, findings, and recommendations from a desk review of the qualitative research conducted within Results-Based Financing programmes (RBF) under the Health Results Innovations Trust Fund (HRITF). The review included 17 studies conducted in Benin, Burundi, Cameroon, DRC, Ethiopia, Haiti, Kenya, Kyrgyzstan, Nigeria, Rwanda, Tajikistan, Tanzania, Zambia, and Zimbabwe. The studies reveal a body of high quality work that is consistent with the conceptual framework of RBF schemes, supported by political will, resources, and research capacity. Strengthening the added value of qualitative inquiry in on-going and future qualitative studies may be enabled by small shifts in thinking and practice, in line with a qualitative research paradigm. First, in order to better ground research in an existing country and system specific context, some interrogation of constructs and posited relationships in the existing conceptual framework for intervention/evaluation may be required. Second, to enable more in-depth and richer data that documents working practices and relations under RBF schemes, training of local researchers should place stronger emphasis on entry to the field, gaining trust, building rapport, and sustaining a dialogue with key informants. Third, smaller, more intensive and focused studies targeting fewer sites and smaller samples - but addressing a wider range of methods and informants within the health system - are likely to yield richer data that can support the understanding of how health workers and managers are responding to schemes, and what impact schemes have on service volumes and outputs.

ABOUT THIS SERIES:

This series is produced by the Health, Nutrition, and Population Global Practice of the World Bank. The papers in this series aim to provide a vehicle for publishing preliminary results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character. For free copies of papers in this series please contact the individual author/s whose name appears on the paper. Enquiries about the series and submissions should be made directly to the Editor Martin Lutalo (mlutalo@worldbank.org) or HNP Advisory Service (healthpop@worldbank.org, tel 202 473-2256).

For more information, see also www.worldbank.org/hnppublications.



1818 H Street, NW
Washington, DC USA 20433

Telephone: 202 473 1000
Facsimile: 202 477 6391
Internet: www.worldbank.org
E-mail: feedback@worldbank.org