

**Measurement, Learning, and
Evaluation Framework for the
Bihar Initiative**

Revised Report

April 18, 2011

Anu Rangarajan
Kimberly Smith
Evan Borkum
So O'Neil
Anna Christensen



MATHEMATICA
Policy Research, Inc.

This page has been left blank for double-sided copying.

Contract Number:
4993/18924(13)

Mathematica Reference Number:
06893.200

Submitted to:
Bill and Melinda Gates Foundation
PO Box 23350
Seattle, WA 98102

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005
Project Director: Anu Rangarajan

**Measurement, Learning, and
Evaluation Framework for the
Bihar Initiative**

Revised Report

April 18, 2011

Anu Rangarajan
Kimberly Smith
Evan Borkum
So O'Neil
Anna Christensen

MATHEMATICA
Policy Research, Inc.

This page has been left blank for double-sided copying.

CONTENTS

A.	Introduction	1
B.	Overview of the Bihar Initiative	2
	1. The Bihar Context	2
	2. Description of the Bihar Initiative and Its Grant Portfolio	3
	3. Objectives of MLE for the Initiative	5
C.	Measurement, Learning, Evaluation (MLE) Approach	7
D.	Results Frameworks	8
	1. Logic Model Development	10
	2. Description of the Initiative-Level Logic Model	10
	3. Identification of the Initiative’s Underlying Theory of Change and Key Assumptions.....	11
E.	Preliminary Learning and Evaluation Questions	13
	1. Identification of Broad Learning and Evaluation Questions	13
	2. Prioritized Learning and Evaluation Questions	16
F.	Measurement Options and Design Parameters.....	19
	1. Design Options.....	19
	2. Sample Size Requirements for Impact Analyses	21
	3. Data Sources	24
G.	Next Steps	26
APPENDIX A: MEASUREMENT FRAMEWORK		A.1
APPENDIX B: LOGIC MODELS FOR THE GRANTS		B.1
APPENDIX C: PRELIMINARY SET OF INDICATORS.....		C.1

This page has been left blank for double-sided copying.

TABLES

1	Underlying Theory of Change and Key Assumptions—Bihar Initiative.....	12
2	Bihar Initiative Learning and Evaluation Questions and Corresponding Measurement Goals and Results Framework Components	17
3	Minimum Detectable Impacts (MDI) and Required Sample Sizes	22
C.1	Preliminary Indicators for Bihar Initiative	C.5

This page has been left blank for double-sided copying.

FIGURES

1	Planned Coverage of Essential Health Delivery and Innovative Solutions in Bihar	4
2	Measurement, Learning, and Evaluation Process	7
3	Logic Model for the Bihar Initiative	9
A.1	Measurement Framework	A.4
B.1	Shaping Demand and Practices to Improve Family Health in Bihar (BBC World Service Trust)	B.3
B.2	Family Health Initiative in Bihar (CARE)	B.5
B.3	Engaging Private Providers to Improve Management of Tuberculosis, Visceral Leshmaniasis, Childhood Pneumonia, and Diarrhea in Bihar (World Health Partners)	B.7

This page has been left blank for double-sided copying.

A. Introduction

The idea that everyone should have the “opportunity to live a healthy, productive life” guides the mission of the Bill & Melinda Gates Foundation. The foundation works to improve global health through a combination of better technology, improved knowledge and practices, and partnerships with local stakeholders. The Global Health Program harnesses advances in science and technology to save lives in poor countries and combats infectious diseases, such as diarrheal diseases, HIV/AIDS, malaria, pneumonia, tuberculosis, and neglected and other infectious diseases. Where proven tools exist, the foundation supports sustainable ways to improve their delivery; where they do not, the foundation invests in research on and development of new interventions, such as vaccines, drugs, and diagnostic technology.

The foundation is committed to reducing maternal, neonatal and child mortality through the sustainable delivery of life-saving tools, technology and approaches. The Global Health Program promotes integrated health solutions for family planning; nutrition; maternal, neonatal, and child health; and vaccine-preventable diseases. It also encourages building the capacity of governments and other development partners to scale up and sustain at scale the delivery of effective interventions. As part of its Global Health Program, the foundation has invested more than \$13 billion since 1994 to investigate new methods to improve health and strategies for successful delivery of proven health-delivery practices, supporting grants in Africa, Asia, Australia, Europe, and North and South America.

The Family Health Initiative in Bihar, India (referred to in this report as the “Bihar Initiative”) is one of the foundation’s flagship programs. It represents a new approach to investing in global health, with the goal of yielding greater impacts on health outcomes and mortality, and accelerating progress toward Millennium Development Goals 4 and 5. In particular, the Bihar Initiative takes an integrated approach to improving reproductive, maternal, neonatal, and child health by leveraging and bundling services and delivery mechanisms from several of the foundation’s Global Health Strategies to improve uptake and coverage across the continuum of family health care. These strategies include Maternal, Neonatal, and Child Health; Family Planning; Nutrition; Vaccine Delivery; Tuberculosis; Enteric and Diarrheal Diseases; Pneumonia; and Neglected and Other Infectious Diseases.

Given the seminal work taking place under the Bihar Initiative, the foundation wants to ensure that the Initiative includes a strong measurement, learning, and evaluation (MLE) component. Specifically, the foundation is interested in understanding program implementation successes and challenges, the cost-effectiveness of innovative and integrated family health solutions, and whether successful solution levers or delivery models can be brought to scale.

To achieve its MLE objectives, the foundation has contracted with Mathematica Policy Research to assist in the development and execution of an MLE plan for the Bihar Initiative. This work will be conducted in two phases: In the first phase, to be completed in June 2011, Mathematica will develop an Initiative-level MLE plan in close collaboration with the foundation, the grantees and other MLE partners. The MLE plan will provide a comprehensive and systematic approach to measuring, monitoring, and evaluating the implementation and effectiveness of the Initiative’s grant portfolio and overall strategy, as well as specific solution levers. It will include prioritized learning and evaluation questions and an approach for answering each question, including the design, outcomes, sample size requirements, data sources, and analysis plan. It will also include a work plan and timeline and will provide an understanding of the roles and responsibilities among grantees and

MLE partners in the execution of the overall MLE plan. The second phase, likely to extend through 2016, will involve implementing the MLE plan and disseminating findings to various stakeholders.

Mathematica will collaborate closely with other MLE partners to implement the MLE plan. Current MLE partners contracted by the foundation for the Bihar Initiative include the Center for Global Health Research (CGHR) at the University of Toronto and the Collaboration for Health Systems Improvement and Impact Evaluation in India (COHESIVE-India).¹ CGHR will lead the effort to identify and analyze appropriate secondary data that can be used to address the Initiative's prioritized learning questions, and COHESIVE-India will conduct a rigorous evaluation of the private sector solutions being implemented by one of the grantees (World Health Partners) funded under the Initiative. In addition to these MLE partners, Mathematica will also identify and work closely with local, Indian partners on primary data collection and on various research activities to be conducted as part of the MLE effort. In both the design and execution phases of the MLE effort, we will also provide technical assistance to grantees, as needed, on their grant-level monitoring and evaluation activities.

This report, referred to as the "MLE Framework" report, describes our approach to developing the MLE plan for the Bihar Initiative. More specifically, we present the process through which we propose to develop the plan, as well as preliminary, key inputs for the plan, including results frameworks, learning questions, and broad evaluation design parameters. These inputs are intended to serve as the basis for in-depth MLE-planning discussions with the foundation and grantees over the coming weeks.

The structure of the report is as follows. The next section contains a brief overview of the Bihar Initiative, including the Initiative's MLE objectives. Section C presents our approach to developing the MLE plan. Subsequent sections of the report present initial steps taken toward generating the MLE plan, including the creation of results frameworks (Section D), the identification of an initial set of learning and evaluation questions (Section E), and specification of key measurement parameters (Section F). Finally, Section G presents next steps toward the development of a detailed and comprehensive MLE plan for the Bihar Initiative.

B. Overview of the Bihar Initiative

1. The Bihar Context

Bihar is one of India's most populous and poorest states, and its health and development indicators point to a reinforcing cycle of poverty and poor health. Bihar's literacy rates are the lowest in the country and its per capita income is less than a quarter of the national average (US \$1,070).² The state also faces continuing public health challenges. Accounting for 8 percent of India's population and 10 percent of its annual births, Bihar contributes to 12 percent of maternal deaths, 12 percent of neonatal deaths, 13 percent of non-fully immunized children, and 15 percent of

¹ COHESIVE-India was formed by researchers from Duke University, Stanford University, University College of London, and the World Bank.

² Total population and literacy rates are available at <http://gov.bih.nic.in/Profile/CensusStats-01.htm>. Per capita income figures are available at <http://pbplanning.gov.in/pdf/Ranking%20of%20States%20Current.pdf>. Per capita income figure is based on 2008-2009 prices.

underweight children.³ Efforts to improve the health situation in Bihar are hampered by health system weaknesses, including gaps in infrastructure and human resources, related inadequacies in the coverage of essential family health interventions, and low levels of knowledge of and demand for appropriate reproductive, maternal, neonatal, and child health services.

Under strong government leadership, the Government of Bihar (GoB) has made major strides in the past several years, improving the overall climate of development in the state and introducing new policies in several spheres including physical infrastructure, education, and health. Despite these strides, the health status of the Bihari population, particularly those residing in rural areas, still needs considerable improvement. Responding to this need, several international donors have made large health sector investments in Bihar in recent years. For example, in 2010, the United Kingdom's Government Department of International Development (DFID) invested 600 crores INR (\$135 million) to reduce maternal and child deaths, under-nutrition, and unwanted pregnancies/fertility through increased scale and functionality of health services, systems strengthening, and greater engagement of non-government actors. In addition, UNICEF has been operating in Bihar for several years to scale-up comprehensive newborn care, strengthen routine immunizations, and address human resource shortages.

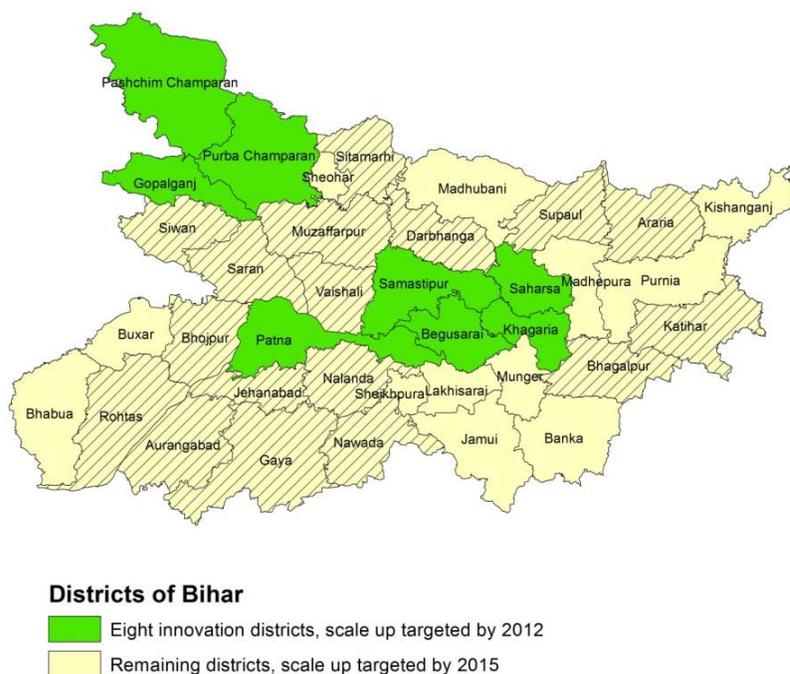
2. Description of the Bihar Initiative and Its Grant Portfolio

The overall goal of the foundation's Bihar Initiative (2010–2015) is to reduce mortality and improve key nutrition and health outcomes by developing innovative and integrated solutions that involve both the public and private sectors. Poor health indicators, high burden of disease, and inadequacies in the supply of and demand for high-impact reproductive, maternal, and child health interventions suggest a need for the broad-based family health approach that the Bihar Initiative has adopted. It addresses both supply- and demand-side barriers to increased uptake, coverage, and quality of family health interventions through a synergistic set of complementary grants working towards a common goal.

More specifically, the Initiative aims to expand the reach, coverage, and quality of (1) essential primary health and nutrition services for infants, children, and women of reproductive age; and (2) diagnostic and disease control services for infectious diseases, including pneumonia, diarrhea, tuberculosis, and visceral leishmaniasis. To achieve these goals, the Bihar Initiative focuses on leveraging and strengthening existing public and private sector delivery platforms. As part of this effort, the foundation has signed a memorandum of cooperation with the GoB that formalizes the GoB's commitment to improving family health indicators and establishes mechanisms for systematic support from the GoB for implementing and scaling-up successful family health solutions.

A key component of the Bihar Initiative is the scaling up of successful supply- and demand-side approaches to the delivery of effective family health interventions. To this end, the Bihar Initiative will focus initially on implementing, testing, and documenting the success of promising solution levers or delivery models in eight innovation districts (Figure 1). Based on the findings from

³ Bill & Melinda Gates Foundation Request for Letter of Inquiry, Family Health Initiative for Bihar (2010-2015).

Figure 1. Planned Coverage of Essential Health Delivery and Innovative Solutions in Bihar

Note: BBC and CARE grants will scale up throughout the state of Bihar by 2015. WHP plans to scale up to 25 districts by 2012; these include the 8 innovation districts and the 17 districts indicated by diagonal lines.

these efforts, the Initiative will then promote and facilitate the replication and scale-up of successful strategies by the GoB and other partners in the remaining 30 (of 38) districts in Bihar.⁴

Under the Bihar Initiative, the foundation will support four complementary grants. Viewed together, these grants reflect the program approach and objectives outlined above. Two of the grants aim to strengthen the supply of essential family health interventions and infectious disease treatments in the public and private sector respectively; one demand-side grant focuses on changing behaviors and norms to support family health; and a community mobilization grant focuses on strengthening grassroots mechanisms for improving the delivery and uptake of family health services. Three of the four grants were awarded in the fall of 2010; the fourth (community mobilization) grant will be awarded in 2011.

The four grants are summarized below:

1. ***Family Health Initiative (FHI) in Bihar.*** The FHI, to be implemented by CARE USA and CARE India, intends to catalyze a dynamic process of development, testing and scale-up of innovative solutions to improve the coverage, quality, and uptake of critical family health services in Bihar. The solutions to be developed and tested can be broadly categorized as strengthening data-driven management; integrating the delivery of family health services; improving tools for frontline workers; and their capabilities and performance; and creating partnerships with private sector family health providers

⁴ One of the grantees funded under the Initiative, World Health Partners, plans to implement in 12 districts in the first year, and to scale up to 25 districts in the second year.

to extend the reach and quality of care. The FHI will collaborate closely with the GoB to encourage support, adoption, and scale-up of solutions that are shown to be cost-effective and have a high impact, with an initial scale-up to eight focus districts and a subsequent scale-up to the entire state.

2. ***Shaping demand and practices to improve family health in Bihar.*** The Shaping Demand and Practices grant, to be implemented by the BBC World Service Trust, aims to increase demand for key family health services and improve family health practices by increasing knowledge, changing attitudes, and shaping social norms. It focuses on developing and testing innovative ways of providing health information through various channels, including mass media, mobile and internet technology, community mobilization, and interpersonal communication from frontline workers. The project intends to build the capacity of GoB to sustain and scale up successful behavior-change communication approaches to influence mortality and health outcomes in the long run.
3. ***Engaging private providers to improve management of tuberculosis, visceral leishmaniasis, childhood pneumonia, and diarrhea.*** This supply-side grant, to be implemented by World Health Partners (WHP), aims to improve the quality of care provided by the private sector for select infectious diseases. The primary objective of the grant is to establish a high-quality, branded private sector health service delivery network (SKY centers) by engaging and training existing private providers to improve detection, diagnosis and treatment of tuberculosis, visceral leishmaniasis, childhood pneumonia, and diarrhea. The strategies for achieving this aim are improving the service delivery system and supply chain for diagnostic tests and treatments; improving the capabilities of private health providers through training and monitoring; stimulating consumer demand for high quality care; creating public-private partnerships; and establishing the sustainability of the private provider network. This project will cover 12 districts in 2011, and expand to 25 districts in 2012.
4. ***Community mobilization.*** This grant has yet to be funded. Its goal will be to create and strengthen community platforms to improve social accountability in the health system and to change health-seeking behaviors and family health practices through increased knowledge and self-efficacy.

Although each of these grants has a different primary focus, they are intended to be complementary and, through synergies across grants, the foundation aims to increase coverage of critical and efficacious interventions and ultimately reduce maternal and child mortality.

3. Objectives of MLE for the Initiative

The overall objective of the Bihar Initiative's MLE effort is to provide credible evidence on the implementation and success of the Initiative's approach, including specific family health solutions being tested, as well as to track progress and inform decision-making by the foundation and grantees throughout the Initiative's lifecycle. Specific measurement, learning, and evaluation objectives for the Initiative include:

- ***Documenting implementation progress, fidelity, successes and challenges, and program costs.*** The MLE effort will include rigorous process evaluations to document and assess achievements and challenges in the implementation of specific solution levers, integrated delivery models, and the overall Initiative. These qualitative evaluations will focus on *how* solutions and models are being implemented and the extent to which

interventions deviate from original plans. They will also be used to identify and document the planning needed for rolling out such interventions, as well as the challenges encountered and how they were resolved. Understanding which family health solutions were implemented and how they were implemented will provide a context for interpreting changes in outcomes that we observe. It will also offer important information and lessons for program improvement during the Initiative's life cycle and for future work. Finally, careful documentation of program costs and key cost components will be critical to understanding the magnitude and scope of the investments made by the foundation and other key stakeholders. It will also be important to document the costs of specific solution levers being implemented by grantees in order to assess the cost-effectiveness of these innovations.

- ***Demonstrating the effectiveness and cost-effectiveness of innovative family health solutions.*** A key component of the MLE effort will be to identify and test the effectiveness of two or three promising solution levers or delivery models being implemented by the grantees. Demonstrating effectiveness will require a rigorous evaluation design that will allow for estimation of the *causal* effect of a given innovation or model—in other words, a design that will allow for attribution of a change in outcomes to a particular intervention. This type of evaluation tends to be more resource-intensive and require a more controlled and limited rollout of an intervention. Therefore, only innovative solutions for which a high degree of certainty about effectiveness is required for replication and scale up decisions will be selected for such testing. The results from these analyses can be used to assess the benefits of the innovations, which in conjunction with cost estimates can be used to derive the cost-benefits or cost-effectiveness of the innovations. The foundation and grantees can use evidence from this MLE component to advocate for adoption of new solution levers by the GoB and the private sector, each of which will play a key role in the scale-up and the long-term sustainability of the innovations being tested under the Bihar Initiative.
- ***Measuring the overall contribution of the Bihar Initiative.*** Initiative-level impacts tend to be longer term goals that are achieved through the combined efforts of several partners, including the foundation, the grantees, the GoB, donors, and communities. As a first step, measurement at the Initiative level will focus on broadly assessing the effects of the foundation's investments on changes in intermediate-term outcomes which are critical to achieving long-term impacts. This will involve comparing outcomes in the eight innovation districts that are expected to be directly affected by the grantees' efforts with outcomes in other selected comparison districts as well as in the rest of the state. Measurement of the overall contribution of the Initiative will focus on a key set of common indicators across grantees, and will not seek to *attribute* all observed changes in outcomes to the foundation's efforts alone. Information obtained from the process analysis on health innovations implemented by the GoB and other donors will provide information on the roles played by the various partners in improving health indicators for families in Bihar.
- ***Identifying factors related to successful scale-up.*** A key assumption of the Bihar Initiative is that integrated solution levers can be brought to scale, and that delivery at scale of high-impact and cost-effective family health services and interventions will significantly reduce maternal, neonatal, and child mortality and morbidity. Given the importance to the Bihar Initiative's success of achieving scale, the MLE effort will include measurement of the scale-up process. This measurement activity will focus on understanding how diffusion of family health solutions occurs and the process of

adoption, uptake, and implementation—initially within the eight priority districts and eventually in the entire state of Bihar. This MLE effort will also involve identifying and documenting best practices related to scale up and diffusion.

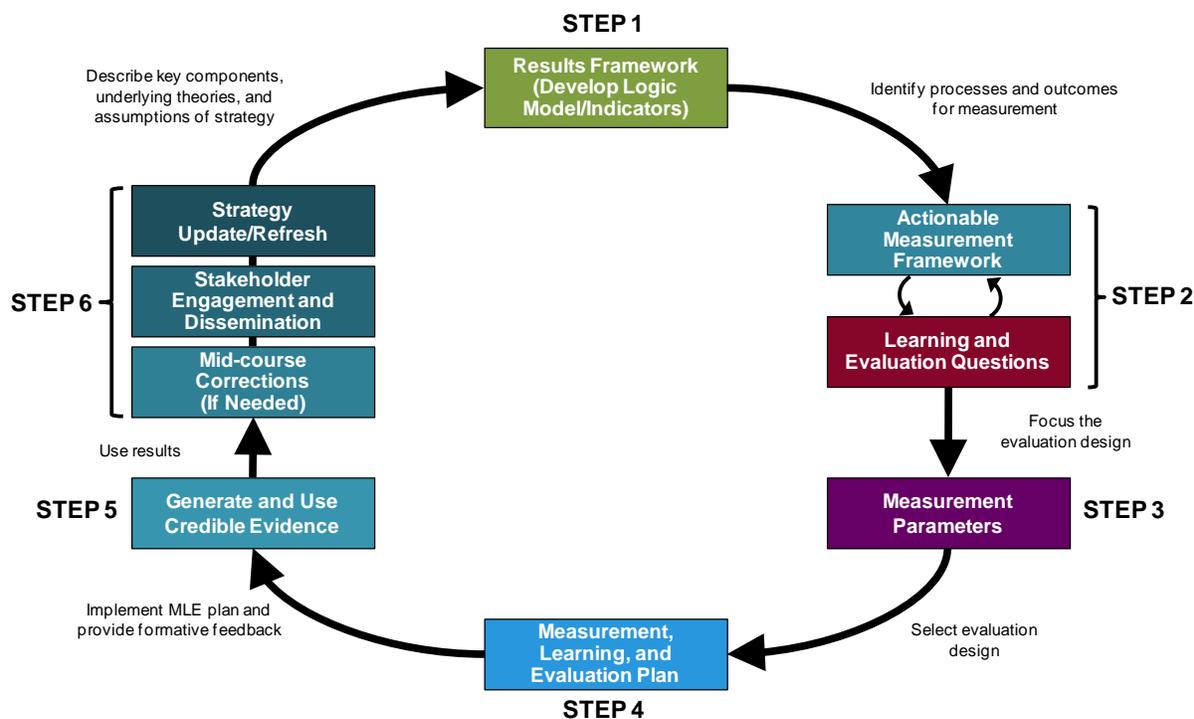
- **Monitoring and tracking process at the grant level.** Monitoring and tracking activities and outputs, as well as outcomes, at the grant level is an essential and core MLE task. This task will be led and executed by grantees, with technical assistance, as needed, from the MLE partners. Each grantee will have their own monitoring and evaluation (M&E) activities, and data from grant-level M&E efforts will be a key input for the implementation and scale-up analyses, and will help inform Initiative-level MLE decisions.
- **Contributing to the field.** An important goal of the foundation is to further the reproductive, maternal, and child health fields through effective dissemination of evidence and lessons learned generated from the Bihar Initiative. Such information can provide critical input for funding and planning decisions by key partners in India and for more widespread efforts to advocate for the replication and expansion of successful family health solution levers.

C. Measurement, Learning, Evaluation (MLE) Approach

Achieving the above measurement objectives will require a comprehensive and systematic MLE approach. This section presents a high-level overview of the MLE process, and how the foundation, grantees, foundation, partners, and external stakeholders can use the MLE results to guide decision making.

Figure 2 shows the MLE process over the Initiative’s lifecycle. The right side of the figure depicts the process through which the MLE plan is created, and the left side shows how the

Figure 2. Measurement, Learning, and Evaluation Process



information generated from the execution of the plan can be used for program improvement, policy and advocacy, and strategic planning within and outside the foundation.

The development of a MLE plan begins with the creation of a results framework (**Step 1**), which is used to identify key processes and outcomes for measurement and initial learning and evaluation questions. A closely related component of the results framework is a core set of indicators that can be used to measure progress toward the achievement of stated outcomes.

Step 2 involves using a measurement framework to guide the identification and prioritization of key learning and evaluation questions by tying questions to the purposes of measurement. Measurement results will be used to inform program improvement and initiative decision making needs, as well as for purposes of advocacy and to further the field. Given the wide range of activities being implemented as part of the Bihar Initiative, a range of questions could be developed to address various aspects of the program. However, measurement is costly and not every question will yield needed, meaningful and actionable information. Having clearly articulated questions that address the Initiative's key measurement needs is an integral component of developing the MLE plan. The measurement framework (described in Appendix A) identifies the goals for and utility of measurement to ensure that MLE efforts are purposeful and focused; it also provides criteria for prioritizing specific learning and evaluation questions.

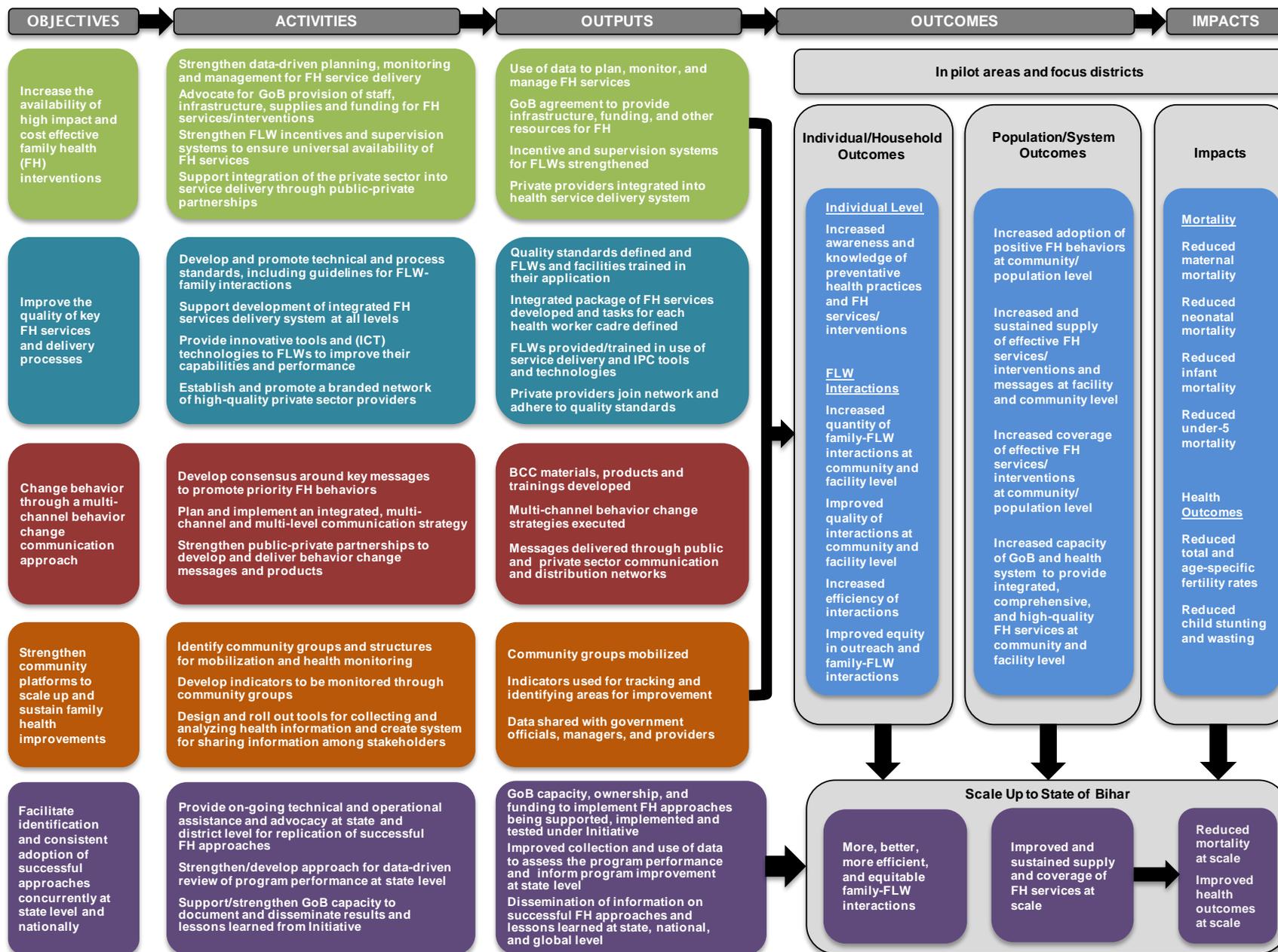
After identifying a prioritized set of learning and evaluation questions, the next step in the MLE process (**Step 3**) is to identify broad evaluation approaches to answering each question. The appropriateness of a particular evaluation approach will depend on the degree of attribution (or level of certainty) required to provide credible evidence on a given question. This step includes the consideration of broad design options, as well as other MLE parameters, such as data needs and sample size requirements. Design options have to be considered in the context of program implementation plans and realities. The selection of the evaluation approach for each question provides the starting point for the development of a detailed evaluation design and forms the basis for the MLE plan (**Step 4**).

The execution of the MLE plan will generate rigorous and credible evidence that can be used to assess whether the underlying assumptions of the strategy hold and if any changes should be made, and to document the implementation and effectiveness of the strategy (**Step 5**). This information can be used by a variety of internal and external stakeholders for various decision making purposes (**Step 6**). For example, the evidence can be used by the grantees and the Initiative to assess progress toward goals and to make mid-course corrections, and to inform the foundation's strategy refresh processes. Similarly, the evidence can be used by the GoB and external stakeholders to guide annual funding and planning decisions, and to advocate for replication and scale up of successful solutions.

D. Results Frameworks

Results frameworks provide a comprehensive and systematic picture of a program and its underlying theory of change by representing the relationships among program activities and how they lead to impacts. They provide a critical foundation for the MLE plan because they describe the activities and outcomes that would need to be measured. Our results frameworks consist of logic models for the Bihar Initiative (Figure 3) and its three approved grants (Appendix B). Ultimately, measurement will require the translation of the concepts from the logic model into specific

Figure 3. Logic Model for the Bihar Initiative



indicators on which information can be obtained to measure progress toward the outputs and outcomes identified in the logic model (see Appendix C).⁵

The Initiative-level logic model describes the objectives of the Initiative, the activities or solutions to be implemented, and the outputs, outcomes, and impacts that might be expected as a result of investments. It facilitates a common understanding of the Initiative's goals and objectives across partners and grounds the MLE plan. Below, we describe our process for developing the logic model, present the draft logic model for the overall Bihar Initiative, and describe the underlying theory of change and key assumptions for the Initiative and the grants under it.

1. Logic Model Development

To help inform the development of the Initiative-level logic model, we first created logic models for each individual grant. The grant-level logic models were informed primarily by the grant proposals submitted to the foundation and should, therefore, be considered preliminary in nature. We plan to update and refine these draft models based on discussions with grantees and the grant-specific 100-day implementation plans. In the current grant-level models, we attempted to standardize terminology across models, where appropriate, for consistency. As a result, the terminology in the logic models may differ slightly from that used in the individual grant proposals.

The Initiative-level model is, in part, an aggregation and reorganization of the individual grant model components. However, it also incorporates additional policy- and advocacy-related objectives at the Initiative level that may not be reflected in the grantees' logic models. In addition to the grant proposals, we also drew on other resources to inform the Initiative-level model, including the foundation's draft log frame for the overall Initiative, foundation slide decks related to the Bihar Initiative, and foundation/partner presentations and materials from the Bihar partners' meetings in December 2010 and March 2011. Discussions with foundation staff were also considered in the creation of the logic models.

2. Description of the Initiative-Level Logic Model

Figure 3 presents the logic model for the overall Bihar Initiative. Reading across the columns from left to right, the objectives of the Initiative are paired with specific activities that will result in observable outputs that contribute to the achievement of outcomes. Outcomes then contribute to impacts. The linkages among inputs, activities, outputs, outcomes, and impacts represent the theory of change and key assumptions that underlie the Initiative's approach. The theory and key assumptions can also be tested as part of the MLE plan as a means of identifying the contribution of Initiative's activities to outcomes.

The individual logic model components can be described as follows:

- **Objectives.** The logic model includes five Initiative-level objectives: (1) to increase the availability of high-impact and cost-effective family health interventions; (2) to improve the quality of key family health services and delivery processes; (3) to change behavior

⁵ Appendix C includes a very preliminary list of Initiative-level indicators compiled from foundation and grantee documents. This list will be refined and selected indicators will be identified for measurement once the logic models are finalized and prioritized learning questions are identified.

through a multi-channel behavior change communication approach; (4) to strengthen community platforms to scale up and sustain family health improvements; and (5) to facilitate identification and consistent adoption of successful approaches concurrently at state level and nationally.

- **Activities.** Activities reflect the inputs being implemented by the grantees as part of the Initiative. An illustration of an activity to increase the availability of high-impact and cost-effective family health interventions is strengthening data-driven planning, monitoring and management for family health service delivery. Similarly, an activity to improve the quality of key family health services and delivery processes is to develop and promote technical and process standards, including guidelines for interactions between families and frontline workers.
- **Outputs.** The activities implemented by the grantees, in collaboration with the GoB, other partners, and foundation, are expected to result in outputs that indicate the direct and tangible execution of the activities. Only indicators that can be categorized as the immediate results from implementation of specific solution levers are considered outputs. Examples of outputs might include the development of behavior change communication materials and tools and the training of frontline worker in the use of information communication technology tools.
- **Outcomes.** If the underlying theory of change and the assumptions of the Initiative are valid, the outputs from the activities are expected to result in improved outcomes. In the logic model, outcomes are presented in three categories.
 - *Individual/household and frontline worker outcomes.* These include outcomes related to awareness, knowledge, and attitudes and to changes in interactions with frontline workers, including the number, quality, efficiency, and equity of interactions.
 - *Population/systems outcomes.* These outcomes demonstrate the reach of the Initiative beyond the individual and interactions level to the facility, community, population, and systems level. They are related to changes in the adoption of preventative health practices; the supply of and demand for family health services at the broader facility, community, and population level; and the capacity of the GoB and the health system at large to provide integrated, comprehensive, and high-quality family services at the community and facility level.
 - *Impacts.* Impact indicators relate to longer-term goals that might require more than 5 to 10 years to manifest or to observe. Impacts relate to key mortality (maternal, neonatal, infant, and under-5), fertility, and nutrition measures.

3. Identification of the Initiative's Underlying Theory of Change and Key Assumptions

After building the Initiative-level logic model, we attempted to identify the major theory of change and assumptions underlying the interconnection or progression from objectives to activities, outputs, and outcomes, drawing on relevant foundation and grantee documents and presentations. Table 1 presents the results of this exercise, which provide important information relevant to identifying key learning and evaluation questions for the Initiative, as the success of the entire effort hinges on the validity of the Initiative's underlying assumptions. For example, effective collaboration between the foundation/grantees and the GoB is essential to the Initiative's success, hence it will be

Table 1. Underlying Theory of Change and Key Assumptions—Bihar Initiative

Initiative/Grant	Underlying Theory of Change/Key Assumptions
Bihar Initiative	<ul style="list-style-type: none"> • Implementation and support of effective solution levers by GoB will contribute significantly to improved family health. In particular, GoB will: <ul style="list-style-type: none"> - Commit infrastructure, funds, and human resources to introduce innovations; new tools and approaches; training and capacity building; and monitoring and supervision of solution levers proven effective by the Initiative - Have the ability and resources to deliver on commitments made in memorandums of cooperation signed with the Foundation - Demonstrate openness and commitment to integrating and partnering with private sector and other players in health care delivery and behavior change communication - Take an active role in coordinating and implementing family health plans and programs - Integrate key messages into packages for frontline workers, school curriculums, National Rural Health Mission, social media communications, women’s groups, and other community groups - Support and use data generated by audits to develop social accountability systems through existing government programs
Family Health Initiative Grant	<ul style="list-style-type: none"> • Increased capacity of GoB will result in improved family health service system, delivery, and behaviors <ul style="list-style-type: none"> - Demand for family health services can be effectively generated using media and other approaches - Frontline workers are open to changes in the status quo and are willing to try new approaches - Adequate equipment and health facilities and functional supply chains for essential services exist - Training for frontline workers and public health center workers can be effectively conducted using existing structures - Continued and committed political and material support from GoB is provided for testing and scale-up
Shaping Demand and Practices to Improve Family Health Grant	<ul style="list-style-type: none"> • Increased awareness and knowledge of family health services and behaviors will result in appropriate care-seeking behavior, utilization, and health behaviors <ul style="list-style-type: none"> - Mobile and mass media penetration grows - Improvement in service provision sufficient to satisfy increase in demand resulting from the grant - GoB recruits and trains frontline workers appropriately; master trainers are effective in training frontline workers - Support from GoB in leveraging resources and encouraging the adoption of mass media activities is provided - Community groups operate effectively and adopt the materials provided - Private sector is interested in collaborating and government policies allow for such collaboration
Engaging Private Providers to Improve Management of Tuberculosis, Visceral Leshmaniasis, Childhood Pneumonia, and Diarrhea Grant	<ul style="list-style-type: none"> • Increased provider training, incentives, and accountability will lead to more comprehensive, effective, and efficient family health services <ul style="list-style-type: none"> - Private providers willing and motivated to make significant investments to improve management of infectious diseases - Medical community does not block engagement of providers - Quality manufacturers willing to produce dispersible amoxicillin and dispersible zinc at low cost - Revised National Tuberculosis Control Programme, National Vector-Borne Disease Control Program budgets sufficient to cover increased payments under incentive schemes for providers - Market dynamics in favor of inappropriate care are surmountable - GoB support for the program is sufficient - Additional revenue streams to support incremental training, communication, and overall management beyond the grant are made available

important to measure the financial and non-financial support provided by the GoB to support the implementation and scale up of family health solutions.

E. Preliminary Learning and Evaluation Questions

This section presents a preliminary list of learning and evaluation questions for the Bihar Initiative. We developed these questions using a three-step process. First, we used the Bihar Initiative’s MLE objectives and the Initiative-level logic model (and the key assumptions underlying it) to inform the development of a broad set of learning and evaluation questions that cut across the individual grants. For each broad question, we then developed an extended set of more specific questions that focus on various aspects of learning at both the Initiative and grant levels. The identification of learning questions related to individual grants was informed primarily by the grant-specific logic models and theories of change. Finally, we used the measurement framework (described in Appendix A) to identify a prioritized set of initial learning and evaluation questions.

The learning and evaluation questions presented in this section should be viewed as preliminary. Over the next two months, we will work with the foundation and grantees to refine and further prioritize questions to include in both the Initiative-level MLE plan and in grant-level MLE planning.

1. Identification of Broad Learning and Evaluation Questions

The learning and evaluation questions for the Bihar Initiative can be grouped into four broad categories that cut across the grants:

1. Which demand- and supply-side approaches to improving reproductive, maternal, neonatal, and child health outcomes were implemented, how were they implemented, and what did they cost?

A strong process or implementation analysis is a critical element of an evaluation. The process analyses of the Bihar Initiative will provide a complete picture of how various solutions levers and activities are being implemented by grantees, whether they are being implemented as planned, and reasons for deviating from the intended plans. The process analyses will be important in helping to understand the facilitators of and barriers to successful implementation of specific approaches to improving family health outcomes. It will also document the context in which the interventions are being implemented, including the roles and activities of key stakeholders such as the GoB, DFID, and UNICEF and assess the extent of collaboration and synergies among grantees and these stakeholders. As part of the process analysis, we will also estimate program and select pilot costs, both overall costs and the costs of major components.

The process and costing studies will provide key input for understanding and interpreting the effects of the overall Initiative and of specific innovations. It will identify potential issues for midcourse corrections, as well as best practices to guide replication of successful interventions in Bihar and elsewhere. They can also help identify the channels through which impacts are manifested so that we can better understand *how* impacts were achieved and at what cost.

Process analyses will use both quantitative and qualitative data, and rely on triangulation of information from various perspectives. Typical data sources for these analyses include project monitoring data, cost data, stakeholder interviews, focus groups, and direct

observations. Information obtained from careful tracking and documentation of program costs will inform the cost-effective analyses.

2. Which innovative solutions or approaches are effective or cost-effective at increasing the coverage of high impact family health and infectious disease interventions?

Each grantee under the Bihar Initiative will implement a range of solution levers to address demand- and/or supply-side barriers to widespread and sustained uptake of high-impact and cost-effective reproductive, maternal, neonatal, and child health interventions. Although these solution levers (or packages of solution levers) may hold great promise, they cannot be scaled up and sustained at scale all at once. Therefore, one objective of the Bihar Initiative is to identify a select subset of successful (cost-effective) solutions and to facilitate the replication and scale-up those solutions to achieve statewide coverage targets.

A key component of the Bihar MLE effort will be to evaluate rigorously the effectiveness of a select number of solutions. Demonstrating effectiveness will require clear attribution of outcomes to the solutions being tested. Generating credible evidence of impact can create a compelling case for replication and scale-up of an innovative solution by GoB, donors, and private-sector partners.

Only innovative solution levers prioritized by grantees as filling a large service gap and that have the potential for scale up are likely to be subject to a rigorous evaluation. Factors affecting the selection of levers for evaluation include the magnitude of expected effects, the strength of existing evidence on effectiveness, the time frame for implementation, key decision-points, costs of implementing the pilots at scale, and whether credible evidence is needed to bring a given solution to scale. Outcome measurement for these impact evaluations will focus on shorter term outcomes that are likely to experience changes (due the solution lever being tested) within a short (two-year) period and which can be easily measured. In addition, these short-terms outcomes should be a key component of the Initiative's underlying theory of change; in other words, achieving those outcomes is considered necessary for the achievement of longer term outcomes. Once impact estimates are generated, information on costs can be combined to assess the cost-benefits or cost-effectiveness of the solution levers.

One of the rigorous evaluations already planned is to assess the effectiveness of an innovative private sector approach to improving infectious disease outcomes being implemented by WHP. As noted earlier, COHESIVE-India will be conducting this random assignment evaluation of the effects of the WHP approach on four targeted infectious diseases. Through further discussions with the grantees and the foundation, we will identify other solution levers to test using a rigorous impact evaluation design, as well as the outcome measures on which such evaluation efforts will focus.

3. Did the Bihar Initiative contribute to intermediate and long-term changes in family health outcomes in the state of Bihar?

The ultimate goal of the Bihar Initiative (as well as the individual grants under it) is to reduce maternal, neonatal, and child mortality and morbidity. Tracking progress toward these long-term goals over the life of the Bihar Initiative is, therefore, a priority for the foundation's MLE efforts in Bihar. An important component of tracking progress will be measuring changes in intermediate outcomes that are hypothesized to lead to (and be a precondition for) long-term impacts. These intermediate outcomes include, among

others, the adoption of positive family health behaviors, access to and utilization and coverage of effective family health interventions, and the quality of public and private sector health service delivery in priority intervention areas.

Any observed changes in intermediate and longer term family health outcomes are likely to be a function of the combined efforts of many partners, including the foundation, the GoB, other donors including DFID and UNICEF, and various public and private sector partners. Because of the large number of health sector partners in Bihar, it will be challenging to *attribute* all observed changes in outcomes in the eight innovation districts or in the state of Bihar solely to the foundation's investments. As a result, the MLE effort for this question will focus on understanding whether the foundation's investments in Bihar contributed to positive changes in outcomes among the foundation's target populations.

Answering questions related to the overall effects or contribution of the Bihar Initiative will rely on existing data sources, supplemented by additional data collected for measurement of select, key outcomes. The first District Annual Health Survey (AHS) which was fielded in 2009-2010 and provides district level estimates, will supply baseline measures for some intermediate outcomes at the district level, including health service coverage as well as information related to longer term impacts.⁶

While the AHS data can provide useful baseline indicators for several key items, there are concerns about relying solely on this data source for the evaluation of the Bihar Initiative for a few reasons. First, the AHS data will not permit measurement of many proximal indicators related to the provision and quality of care at the household, community, and facility levels. These proximal indicators are more likely to be affected by grant activities implemented within the short timeframe within which scale up decisions need to be made. Second, there is uncertainty about whether household level AHS data will be released for research purposes, which would severely limit its utility for the MLE analysis. Third, there is uncertainty about the timing of and states to be included in the next round of AHS data collection, making it highly risky to rely solely on this source as the primary data source for the MLE effort.

For these reasons, we believe that collection of primary data (on a limited scale) is essential to achieve the MLE objectives, and to provide timely and useful information for decision-making. We propose that baseline data—including surveys of households, frontline workers and facilities—be collected in the eight innovation districts and a set of comparison districts. These data will focus primarily on more proximal outcomes and populations targeted by the Bihar Initiative. A mid-line follow-up should be conducted within the timeframe of 12 to 18 months after baseline, which can be used to assess the effects of grantees' activities before scale up occurs and to inform decisions related to scale up. Finally, we recommend an end-line survey be targeted for the period approximately towards the end of the Initiative which will help assess the longer-term impacts of these interventions.

⁶ The third District-Level Household Survey (DLHS), which was conducted in 2007-2008 includes many of the same measures as the AHS, can provide additional pre-intervention data points for potential trend analysis of intermediate outcome measures.

Data from the follow-up AHS, if conducted in Bihar and in the relevant timeframe for the MLE effort, will complement the new survey data and focus on the analysis of intermediate outcomes and impacts. With respect to measurement of ultimate impacts, accurate measurement of maternal, neonatal, and under-five child mortality can be challenging due to weak registration systems, the relatively low frequency with which mortality (particularly maternal mortality) occurs, and the misclassification/misreporting of the causes of mortality in survey data. CGHR has extensive experience working on these indicators and assessing their quality, and will take the lead on the identification and use of existing data sources—including AHS, DLHS, and any other sources, to assess the Initiative’s contribution to a broad set of coverage indicators and to impacts including mortality change. We will also explore with CGHR the extent to which existing data sources, including household surveys and disease surveillance data files, can be used to measure changes in priority reproductive, maternal, neonatal, and child health outcomes, including child stunting and wasting and infectious disease morbidity.

4. What integrated family health solutions can be brought to scale successfully in the eight priority districts in Bihar and statewide?

A key assumption underlying the Bihar Initiative is that some of the effective integrated family health solution levers being implemented by grantees in collaboration with various partners can be scaled up, and that through delivery at scale of high-impact and cost-effective family health interventions, reductions in mortality, morbidity, and fertility rates will occur. Given the importance of achieving scale to the success of the Bihar Initiative, a key component of the MLE effort will be to measure whether scale-up is happening and to identify best practices related to scale-up.

The Initiative envisions a two-phase scale-up process for most solution levers or delivery models. The initial phase involves the scaling up of essential interventions and successful solution levers in the eight innovation districts within the first two years of the Initiative. The second phase involves scale-up to the entire state of Bihar. However, some interventions (for example, the mass media behavior change communication approaches being implemented by the Shaping Demand and Practice grant) will be implemented at scale immediately.

Expanding the target population served by a pilot to a broader population level involves the development and support of partnerships and the establishment of an enabling environment. Therefore, the measurement of scale-up will involve measurement of the spread of family health innovations in the public and private sectors, mechanisms through which diffusion and dissemination occurred, uptake of key interventions being delivered through innovative solution levers, and key enabling and inhibiting factors for scale-up.

2. Prioritized Learning and Evaluation Questions

For each of the broad questions listed above, we identified more specific learning and evaluation questions (Table 2). Although some of the specific questions cut across grants, others are necessarily grant-specific, as some grants specialize in particular types of solution levers or delivery models. In Table 2, we also indicate the alignment between the learning and evaluation questions and the measurement goals and results framework.

Table 2. Bihar Initiative Learning and Evaluation Questions and Corresponding Measurement Goals and Results Framework Components

Learning and Evaluation Questions	Measurement Purpose/ Decisions to Be Informed	Corresponding Results Framework Outcomes and Impacts
<p>I. Which demand- and supply-side approaches to improving reproductive, maternal, neonatal, and child health outcomes were implemented, how were they implemented, and what did they cost?</p> <p><i>Was there an increase in data use for decision making by frontline workers and managers? What were the barriers to such use?</i></p> <p><i>Which support and incentive mechanisms for frontline workers were implemented successfully? What obstacles did frontline workers face in benefiting from these mechanisms?</i></p> <p><i>Were private-public partnerships successfully formed? Was the Initiative effective in expanding and sustaining the network of private providers?</i></p> <p><i>Did community groups' focus on family health issues increase? Did they successfully adopt and use BCC materials?</i></p> <p><i>What key components of the SKY network (telemedicine, supply chain, provider monitoring systems, and incentives) were successfully implemented? How did these components interact with one another?</i></p> <p><i>To what extent did the grantees integrate their efforts? How did they collaborate, coordinate and support each other during the course of the Initiative? What were the benefits of integration?</i></p> <p><i>What were the costs of providing these services? What were the foundation's costs and how much did GoB contribute? How were the costs broadly allocated across broad program components?</i></p>	<ul style="list-style-type: none"> • Document implementation • Inform grantees' and foundation's decisions related to need for course corrections and program improvement <ul style="list-style-type: none"> – Provide guidelines and criteria for replication • Monitor and track progress <ul style="list-style-type: none"> – Monitor activities and outputs at grant level – Provide information for midcourse corrections • Inform the field <ul style="list-style-type: none"> – Understand successes and challenges of implementation – Understand interaction of grantees under a single initiative 	<ul style="list-style-type: none"> • Increased analysis and use of data for decision-making • Increased supply and utilization of high-quality comprehensive family health services • Increased private sector participation in the Initiative through public-private partnerships • Increased mobilization of community bodies to improve family health • Increase in coverage of infectious disease treatments through high-quality private provider networks
<p>II. Which solution categories or approaches are effective or cost-effective in increasing the coverage of high-impact family health and infectious disease interventions?</p> <p><i>Were incentive schemes for frontline workers effective in increasing quality provision of bundled family health services? What type/size of incentive is most efficient?</i></p> <p><i>Were support mechanisms for frontline workers effective in increasing their motivation and performance? What was the relative contribution of different support mechanisms such as ICT tools and increased supervision?</i></p> <p><i>Did the development of quality of care protocols lead to an improvement in the quality of services rendered by Frontline workers? Did they improve client satisfaction?</i></p> <p><i>Did incentive schemes based on the recognition of excellence among frontline workers, facility teams and private providers improve the quality of care?</i></p>	<ul style="list-style-type: none"> • Demonstrate effectiveness <ul style="list-style-type: none"> – Provide highly credible evidence on effectiveness of selected levers • Inform decision making related to identifying cost-effective solution levers that can be adopted by GoB and scaled up across the state • Inform the field <ul style="list-style-type: none"> – Disseminate results widely to encourage replication 	<ul style="list-style-type: none"> • Increased frontline worker capabilities and motivation for providing key family health services • Increased provision of bundled family health services by frontline workers • Improved frontline worker knowledge and skills around quality standards • Improved adherence to quality standards by Frontline workers

Table 2 (continued)

Learning and Evaluation Questions	Measurement Purpose/ Decisions to Be Informed	Corresponding Results Framework Outcomes and Impacts
<p>III. Did the Bihar Initiative contribute to intermediate and long-term changes in family health outcomes?</p> <p><i>How did family health outcomes change over the course of the Initiative? How did changes in outcomes in the eight innovation districts compare to those in other similar districts? The rest of the state?</i></p> <p><i>Was there an increase in the coverage of essential family health services? Was there an increase in the adoption of preventive family health behaviors? Which types of behaviors were adopted?</i></p> <p><i>Was there an increase in the quality of family health services provided by frontline workers and public and private health facilities?</i></p> <p><i>Can a network of high-quality private providers reduce the incidence and prevalence of infectious diseases?</i></p> <p><i>Do integrated supply and demand side intervention models improve health outcomes relative to demand side models alone? Relative to supply side alone?</i></p> <p><i>Did any of the changes in coverage, uptake and service quality differ for traditionally underserved groups (as defined by poverty or caste)?</i></p>	<ul style="list-style-type: none"> • Measure overall contribution <ul style="list-style-type: none"> – Document changes in long-term outcomes – Measure contribution to changes in long-term outcomes by focusing on intermediate outcomes – Assess the effects of the Initiative's efforts in improving shorter term outcomes • Inform decision making about scale up • Inform the field 	<ul style="list-style-type: none"> • Increased adoption of positive family health and reproductive behaviors • Increased coverage of appropriate family health services and treatments for infectious diseases: <ul style="list-style-type: none"> – Increased utilization – Increased public and private sector supply – Improved equity in coverage • Improved quality of care provided by frontline workers, public facilities and private providers • Improved family health outcomes (mortality, fertility rates and child stunting and wasting) • Reduced infectious disease morbidity
<p>IV. Which integrated family health solutions can be brought to scale successfully in the eight innovation districts in Bihar and statewide?</p> <p><i>To what extent did scale-up occur? How many other districts adopted the identified effective solutions?</i></p> <p><i>What were the key factors behind the successful scale-up of particular solution levers? What were the key barriers that prevented scale-up of others?</i></p> <p><i>How did scale-up occur in practice? Was there spontaneous adoption by non-focus districts? How did scale up in the 8 focus districts differ from scale-up across the entire state?</i></p> <p><i>Did government and private sector support (financial and non-financial) for innovative solutions increase over time? How did the relative roles of different actors change over time? What role did each play in achieving scale-up?</i></p>	<ul style="list-style-type: none"> • Measure scale-up • Inform foundation decision making on carrying forward this model to other states and which components of the model to replicate in those states • Inform the field <ul style="list-style-type: none"> – Understand successes and challenges of scale-up 	<ul style="list-style-type: none"> • Increased financial and nonfinancial support by GoB and private sector • Increased stewardship of family health initiatives by GoB • Increased capacity of GoB to implement and sustain solution levers at scale • Adoption of successful approaches at state level

F. Measurement Options and Design Parameters

Each of the broad learning and evaluation questions proposed in the previous section will require a different methodological approach to measurement, given the nature of the question and the desired degree of attribution (or rigor). Identifying specific measurement approaches to include in the MLE plan will involve extensive discussions with grantees and the foundation, both to refine and prioritize the learning and evaluation questions presented in the previous section and to determine the appropriateness and feasibility of various evaluation designs. The final designs elaborated in the MLE plan will be a function of several factors, including level of attribution needed to answer a given question adequately, grantees' implementation plans, cost, and other resources needed for execution of a given design.

This section describes what we envision as key components (or parameters) of the measurement approach that the MLE plan will elaborate. These components include (1) the appropriate design option to address the study question, (2) the sample size requirements needed for impact analysis, and (3) the potential data sources from which we can obtain key outcomes and other indicators for measurement.

1. Design Options

We expect that answering each of the learning and evaluation questions included in the final MLE plan will, to varying degrees, involve a mixed-methods approach. However, we can broadly classify the general measurement approaches that we will consider into the following types:

- **Impact analysis using a randomized design.** This type of analysis is most appropriate when a high level of attribution is desired. Measuring program impacts typically requires comparing the outcomes for a group exposed to an intervention with the outcomes they would have experienced in the absence of the intervention. The latter, called the *counterfactual*, is typically not observed. Therefore, it is important to create a credible counterfactual to ensure that observed changes can reasonably be attributed to the intervention and not to other factors.

One approach to creating a strong counterfactual is to use a randomized or experimental design, in which individuals or groups (for example, frontline workers or villages) are randomly assigned to either a treatment group that has access to the intervention or a control group that does not (at least for some time). This allocation mechanism ensures the creation of two groups that are equivalent at baseline, with only one group having the opportunity to receive the intervention. Thus, any observed differences in outcomes over time for the two groups can be reliably attributed—with a known degree of certainty—to the effects of the interventions.

Given the high level of attribution required to provide credible evidence on the effectiveness of select innovative solutions being piloted, we propose using a randomized design, where feasible, to address some of the specific questions that will be prioritized under **Question 2** in the previous section (*Which solution levers are effective at increasing the coverage of high impact and cost-effective family health and infectious disease interventions?*).

- **Impact analysis using a comparison group design.** When random assignment is not feasible, we can use quasi-experimental designs to measure impacts of a particular solution lever or innovative model; in a quasi-experimental design, a strong, credible

comparison group is established. The comparison group should include individuals, communities, or geographic units that did not have the opportunity to receive the intervention, but whose behavior can reasonably be assumed to mimic the behavior that would have been observed among participants in the absence of the intervention. With a matched comparison group design (or difference-in-difference approach), identifying a credible comparison group often requires statistical matching on observed characteristics, as well as statistical techniques to control for any existing differences that could confound interpretation of the impact. When the only comparison groups available are weak or nonmatched, the ability to attribute impacts credibly is diminished.

We will consider a matched comparison design for some specific questions under **Question 2**, when randomization is not feasible. A comparison group design may also be appropriate for measuring the overall effects or the contribution of the Bihar Initiative (**Question 3**), depending on the existence of a valid comparison group (for example, a similar district, group of districts, or state).

- **Pre-post or trend analysis.** This type of analysis examines outcomes of interest at one or more points before implementation of a solution lever/model and at one or more points in time after implementation to assess whether outcomes change as expected. An examination of trends over time does not allow for attribution, but significant changes in relevant outcomes before and after implementation of the intervention can signal possible connections between outcomes and the interventions.

We will consider a pre-post or trend analysis for **Question 3** (measuring the overall contribution of the Initiative) if a valid comparison group is not available.

- **Process and costs analyses.** As noted earlier, we recommend a strong qualitative analysis to learn about program implementation and to inform the impact analyses. Qualitative analysis is appropriate for addressing questions about best practices, facilitators and barriers to implementation in different contexts, and to understand the process of scale-up. This analysis generally relies on triangulation of information from various perspectives using an assortment of qualitative data collection and analysis techniques, as well as analysis of project management and administrative data. Although grantees may conduct some of their own implementation analysis, a strong external process analysis will be valuable for documenting the implementation process, as well as the successes achieved and challenges faced. This analysis will also involve creating a framework for collecting cost data for the overall Initiative and for the implementation of select solution levers. We will employ strong qualitative methods, supplemented by programmatic, cost, and management information systems data, to address **Questions 1 and 4**.
- **Diffusion and network analysis.** Although rigorous documentation and analysis of implementation-related data can be used to measure whether scale-up is happening and best practices related to scale-up, other analytic techniques—including diffusion and network analysis—provide additional insight into the mechanisms and processes underlying scale-up. This more targeted analysis can be used to assess the pathways or mechanisms through which family health innovations are spread, and the extent and rapidity of scale-up through these various pathways or mechanisms. Therefore, we will consider this type of analysis for **Question 4**.

2. Sample Size Requirements for Impact Analyses

The design for questions related to demonstrating effectiveness of particular solution levers or approaches will need to pay close attention to issues of statistical power and the sample sizes needed to detect desired effects. Sample size needs for analysis (using either new or existing data sources) is a critical component of any impact analysis as it determines the statistical precision of the impact estimates that can be measured. There is an inherent tradeoff in increasing sample size: a larger sample leads to more precise estimates but increases the costs of data collection (if extant data sources are not adequate). Particularly in the context of testing pilot interventions, the sample size requirements have implications for the number of individuals or places that an intervention covers. In this section, we provide preliminary estimates of sample size requirements for a randomized or comparison group evaluation design in the Bihar context.

As noted earlier, in both randomized and comparison group designs, we will need to have interventions or solutions implemented for some individuals, population subgroups, or geographic units, and not for others (which will represent the counterfactual). Generally, statistical power is much greater if random assignment is at the individual level, but given that most interventions will be implemented at the community level, the relevant unit of assignment will likely be villages (or village clusters), sub-centers, or the PHC (or block).⁷ For example, an incentive scheme for frontline workers or campaigns to improve knowledge and behaviors could be implemented or piloted initially in some sub-centers (or village clusters) and not in others. Other interventions, such as improving the supervision structure may be ideally implemented at the PHC/block level. We present calculations for a randomized or matched comparison design in which an intervention (or treatment) is implemented at the block, sub-center or village level within the eight focus districts. Because residents living in the same village, sub-center or block are likely to face similar conditions, as well as unobserved random shocks (such as an infectious diseases outbreak), it is important to factor in the correlation of households within the same community when determining the sample size needed to estimate effects of a given size.

Table 3 shows the minimum detectable impacts (MDIs)—which represent the smallest impact that can be statistically distinguished from a zero impact—for various sample sizes.⁸ We report the MDIs for outcomes that have a 30 or 50 percent prevalence at baseline. Examples of baseline prevalence for key outcomes of interest in the eight focus districts include: fully immunized child (40 percent); having three antenatal care visits during pregnancy (27 percent); and delivery at a health facility (32 percent).⁹

⁷ Some interventions, particularly related to frontline workers, may be tested at the individual level. However, with individual random assignment, it will be important to ensure minimal spillovers/crossovers between the treatment and control/comparison groups; in other words, if treatment group workers for example talk about or share their tools and practices with control group members (which is likely if they are in the same location), true intervention effects will not be detectable (even though such spillover may be desirable from a programmatic perspective).

⁸ We calculate the power and statistical significance required at the conventional levels of 80 percent and 5 percent, respectively. The intraclass correlation (ICC) accounts for the correlation in outcomes between individuals in the same cluster; we assume a conservative value of 0.10. The extent to which regression methods that control for baseline characteristics can reduce variability in outcomes; we assume an R^2 of 0.3 both within and across clusters.

⁹ Data provided in foundation documents.

Table 3. Minimum Detectable Impacts (MDI) and Required Sample Sizes

Unit of Assignment	Number of Villages or Sub-centers or Blocks (Total)	Number of Respondents per Village or Sub-Center or Block	Number of Respondents (Total)	One Intervention		Two Interventions		Three Interventions	
				Baseline 30%	Baseline 50%	Baseline 30%	Baseline 50%	Baseline 30%	Baseline 50%
				MDI (%)	MDI (%)	MDI (%)	MDI (%)	MDI (%)	MDI (%)
Village/ Sub-center	360	15	5,400	3.7	4.0	4.5	4.9	5.2	5.6
	360	25	9,000	3.6	4.0	4.4	4.8	5.1	5.6
	360	50	18,000	3.6	3.9	4.4	4.8	5.1	5.6
Village/ Sub-center	240	15	3,600	4.5	4.9	5.5	6.0	6.3	6.9
	240	25	6,000	4.4	4.8	5.4	5.9	6.3	6.8
	240	50	12,000	4.4	4.8	5.4	5.9	6.2	6.8
Village, Sub-center, or Block	120	15	1,800	6.3	6.9	7.7	8.5	8.9	9.8
	120	25	3,000	6.3	6.8	7.7	8.4	8.9	9.7
	120	50	6,000	6.2	6.8	7.6	8.3	8.8	9.6
Village, Sub-center, or Block	80	15	1,200	7.7	8.5	9.5	10.4	11.0	12.0
	80	25	2,000	7.7	8.4	9.4	10.3	10.9	11.9
	80	50	4,000	7.6	8.3	9.4	10.2	10.8	11.8
Village, Sub-center, or Block	60	15	900	8.9	9.8	11.0	12.0	12.7	13.8
	60	25	1,500	8.9	9.7	10.9	11.9	12.6	13.7
	60	50	3,000	8.8	9.6	10.8	11.8	12.5	13.6

Notes: Minimum detectable impacts (MDIs) assume an R^2 of 0.3 (within and across clusters) and an intraclass correlation (ICC) of 0.1. The calculations assume a two-tailed test with 80 percent power and 5 percent significance. The total sample is assumed to be divided equally between the control group and the various treatment groups and the MDIs are for the comparison between each treatment group and the control group.

The first two columns of the table note the unit of assignment (village or sub-center or block) in which an intervention may be implemented and the number of units in which the intervention will be implemented. The third and the fourth columns, respectively, show the number of individuals (such as families or frontline workers) responding to a survey in that unit and the total sample size (which is the product of the number of unit and the number of individuals responding per unit).¹⁰

Table 3 presents the MDIs for three options for impact analyses, with one, two, or three treatments being tested. In the most basic impact analysis design, with one treatment group, one intervention or a specific solution lever (or a combination of levers) is tested in certain locations and compared with the outcomes of the control group. However, because we might be interested in assessing the relative effects of different solution levers, we also present the MDIs for the two treatment groups versus a control group, and for three treatment groups versus a control group. In the table, the calculations assume that the total sample (the number of villages or sub-centers or blocks where an intervention is implemented) is divided equally into each treatment group and

¹⁰ It is important to note that the number of respondents indicated in this table represent the sample size needs for the analysis; that is, it represents the sample from which we will obtain data on key outcomes (for example through a household survey or through surveys of frontline workers). These individuals will typically be randomly sampled from the target population in these selected communities to ensure that the impact estimates are representative of the target population for the intervention in these communities. Therefore the “penetration” or “coverage” rates within the treatment clusters need to be high enough that the randomly sampled individuals from the target population are actually exposed to the treatment. In other words, the intervention will target the entire unit (village or sub-center or block) but we will be sampling only a few individuals from that unit.

control group. The MDIs shown are for the comparison between each treatment group and the control group.

The MDIs reported in the table indicate the magnitude of changes we would have to see in a particular outcome to identify interventions that show statistically significant impacts. For example, an MDI of 11 for an outcome with a baseline value of 30 percent implies that, for a given intervention, we would have to see a change in the outcome from 30 percent at baseline to at least 41 percent by the time of the follow-up data collection (a 37 percent improvement in the outcome during that time).

The MDIs presented in Table 3 are only preliminary calculations and will be refined based on updated information on key parameters to be measured and the level at which solution levers are implemented. Nevertheless, several key points emerge from these calculations.

- **The greater the number of units in which an intervention is implemented, the higher the chances of detecting impacts.** Focusing on the MDIs for just one intervention, and for a variable with a mean level of around 30 percent, we note the following: With a sample of 240 villages or sub-centers (120 treatment and 120 control), and surveying 15 target population families per unit, we would be able to detect a significant impact of 4.5 percent change. However, with 120 sub-centers or villages (60 treatment and 60 control), and surveying the same number of target families per unit, we would be able to detect significant impacts only if the true impact was 6.3 or larger. With only 60 villages or sub-centers (30 treatment and 30 control), the MDI increases to 8.9.
- **Increasing the sample size within a unit has little effect on the MDI.** As Table 3 shows, increasing the number of sample families within a unit from 15 to 50 **does not** alter the MDI very much. This is because there are likely to be village- or sub-center- or block-specific effects; in other words, individuals in a village or sub-center or block are more likely to behave like one another or experience the same shocks, and cannot be treated as completely independent of one another. Although we will need a minimum number of households per unit, increasing the sample size does not add as much to power as would sampling the same number of individuals in more clusters.
- **MDIs increase when more interventions are tested using the same number of units.** For example, with a sample of 120 village clusters, for an outcome with a baseline value of 30 percent, we could detect 6.3 percentage points change with one treatment only, 7.7 percentage points change with two treatments, and 8.9 percentage point change with three treatments.¹¹

Because sample sizes within a unit have very little impact on the MDIs, whereas the number of units has a large impact, it would be ideal to have larger numbers of units in which an intervention is being tested. This is particularly true when more interventions are being tested against each other.

¹¹ Baseline prevalence also has some effect on the MDI that can be detected, with smaller MDI for outcomes that are closer to the tails of the distribution than those outcomes with 50 percent prevalence at baseline. This also suggests it is important to obtain accurate information on baseline rates before determining the required sample size.

Ultimately, the choice of the number of units (villages or sub-centers or blocks) needed to test a specific intervention will depend on several factors. The first factor is the change in outcome that is anticipated as a result of the intervention (within the time frame of the data collection or data availability). If there is an a priori expectation that a particular intervention will result in a very large change in an outcome within a short period, fewer treatment units might be adequate to detect impacts than if only small changes are anticipated within the observation window. Second, in addition to what changes in outcomes might be expected, it will be important to ensure that the measured changes are meaningful. For instance, there is probably little value in trying to obtain very large samples sizes to detect very small differences that might not materially improve families' health. Hence, in determining which pilot solutions to test for effectiveness, we will have to identify, with the grantees, the specific interventions or solution levers they expect to change or improve key outcomes the most in a relatively short time.

3. Data Sources

We expect to draw on a combination of existing and new data sources to obtain information on relevant measures needed to answer the learning and evaluation questions. Data will be needed to measure key outcomes as well as to identify factors that might be related to observed outcomes, such as individual and household socio-demographic characteristics, that can be used in the analysis. We will also need data on community characteristics and the availability of existing health facilities. Finally, the answers to questions related to implementation will require a broad range of quantitative and qualitative data, including administrative data, focus group and interview data, and data from direct observation, most of which will require new data collection.

We will identify specific gaps in extant data and new data collection needs as we finalize the learning and evaluation questions and the associated outcomes needed for measurement. This section provides a broad overview of expected data sources/needs.

Extant data. Some data (primarily household survey data) will already be available through well-established and ongoing data collection efforts, including the DLHS, AHS, and NFHS. To the extent possible, the evaluation will draw on these existing datasets, which include several indicators of interest. These data will be particularly relevant to track the contribution of the Initiative to improving family health outcomes in Bihar. The information contained in these data may allow for measurement of maternal, neonatal, and child mortality and some indicators of coverage, as well as the adoption of relevant behaviors service utilization.

New data to be collected. While we plan to take advantage of extant data sources to the extent possible, these sources will not include the full set of measures needed to answer all of the MLE questions. In particular, for rigorous evaluations of specific solution levers tested in select locations, we will need baseline surveys in those areas to establish current prevalence levels and to ensure baseline equivalence of the treatment and comparison groups. Similarly, in order to assess the effects or contribution of the overall Initiative on shorter term outcomes (and some intermediate outcomes), we will need representative information for the eight districts as well as a set of carefully selected comparison districts. New data to be collected could include:

- **Additional household surveys targeting treatment and comparison areas.** The types of data that would be collected through household surveys will be targeted to the interventions being tested and key outcomes they would expect to influence. These data could include knowledge and attitudes related to reproductive, pregnancy-related and neonatal care, the demand for or the use of existing community- and facility-based

maternal, child, and neonatal health (MNCH) and reproductive health (RH) services, the number, nature, types, and quality of interactions between households and frontline health workers, and adoption of preventive MNCH and RH behaviors. Household cultural practices and access to existing family health services in the community will also provide important information on mediators or other factors influencing outcomes, while information on household demographic and socio-economic characteristics will be important for indicators relating to equity.

- **Frontline workers surveys.** Surveys with frontline health workers will assess their levels of capability, motivation, and performance. Information from frontline workers will also include the number of interactions with households and the types of services that they provide during these interactions. Information will also be collected on the types of training they received, the skills they possess, their tenure in their current role, as well as background demographic and socioeconomic characteristics. Frontline workers may also provide valuable information on the difficulty or ease with which households can access existing family health services, as well as perceptions of the level of knowledge that households have in terms of reproductive health, and pregnancy and newborn-related care and practices.
- **Facility surveys.** Facility surveys can provide useful information on outcomes and/or data on measures that can be used as control variables in analytic models. The types of information collected from facilities could include information on medical infrastructure, availability of medical supplies and emergency care services, and the quality of care provided.
- **In-depth interviews with stakeholders.** In-depth or semi-structured interviews with key stakeholders can provide insights into how the interventions are being implemented, and what is working well and what is not. These interviews may also give insights regarding the dynamics of scale-up. In-depth interviews may be conducted with a variety of stakeholders, including program implementers, government staff and local health authorities, community leaders, health workers, and private sector partners.
- **Focus groups.** Focus groups may be conducted with households, frontline workers, health facility staff, or other groups of stakeholders to obtain information related to various types of measures described earlier. The advantages of focus groups are that they draw out group dynamics and provide a means to capturing the perceptions, voices, and personal experiences that can provide rich insight into the interventions but cannot be captured adequately by survey data.
- **Field visits and observations.** Field visits to the local communities, visits with local officials, and observations of facilities will all help provide a sense of the context and the communities in which the interventions are operating. Observations of frontline worker interactions at clinics and in households can provide measures of the quality of frontline worker performance and the knowledge and capabilities workers bring.

In addition to these data, we will also draw on information that may be collected for administrative or other tracking purposes:

- **Grantee tracking data.** Grantees may have management information systems (MIS) or other information reporting systems that will provide information on the extent to which the projects were implemented, the number of trainings conducted for frontline workers,

and other such data. While the extent to which such information is collected will vary by grantee, it will provide useful information on program implementation.

- **Other administrative data.** Health management information systems (HMIS), administrative records, and other data from the government will provide useful information on program implementation and scale-up, particularly if the data quality is good, or is being improved as intended in the Bihar Initiative. Examples include health MIS data, district plans and expenditures, internet and mobile technology usage, program monitoring and financial reports, frontline worker training reports, and sales and distribution data.
- **Disease surveillance data.** Data from various disease surveillance and control programs may be used to identify changes and patterns in infectious disease incidence, prevalence, and mortality. Such programs include the National AIDS Control Organization, the Revised National TB Control Programme, and the National Vector-Borne Disease Control Programme.

G. Next Steps

The next steps toward the development of a detailed and comprehensive MLE plan for the Bihar Initiative include the following, each of which will be undertaken in close collaboration with the foundation:

- Refine and finalize the logic models to ensure they reflect the grantees current implementation strategies and plans.
- Refine and prioritize the learning and evaluation questions to ensure that they capture the key learning objectives of the grantees and the foundation and can be addressed through the Initiative's MLE component; work with foundation staff to select a core set of priority questions on which the MLE plan will focus.
- Develop measurement and evaluation approaches to answering each of the core learning questions that are aligned with grantees' implementation strategies, timelines and realities on the ground.
- Present a finalized MLE plan in June 2011, that will include detailed evaluation designs and plans, including data collection and analyses, as well as a work plan for the overall MLE effort.

Input from the foundation, grantees, and other MLE partners will ensure that the final MLE plan is relevant, feasible, and realistic. The plan will be the overarching guide to measurement for the Initiative; therefore, it is critical that stakeholders participate in its development and that the plan aligns with their learning objectives and implementation plans. While a clearly articulated plan at the outset provides an essential roadmap for planning purposes, we are very aware that the MLE plan may need to be modified over the course of the Bihar Initiative's lifecycle in response to changing conditions and lessons learned on the ground.

APPENDIX A
MEASUREMENT FRAMEWORK

This page has been left blank for double-sided copying.

Measurement Framework

1. What Is a Measurement Framework?

A measurement framework is a systematic process or approach used to ensure that MLE efforts are relevant, meaningful, and focused. It guides the identification of the purposes, priorities, and methods for measurement, including the prioritization of key learning and evaluation questions. It also establishes criteria that can be used to ensure that measurement is feasible and that MLE results are actionable. Ultimately, the measurement framework guides the development of key inputs for the MLE plan.

Figure A.1 presents an illustrative measurement framework. It starts with the broad purposes of measurement (“Why measure”), and then identifies measurement utility (“How will measurement results be used”). The purposes and utility of measurement inform the identification of key learning and evaluation questions (“What to measure”) based on logic models and the theory of change underlying them. Criteria specified in the framework are then used to prioritize the learning and evaluation questions. Such criteria include: (a) the feasibility of measurement, (b) whether the information generated through measurement can be acted upon (by the foundation, grantees, or external partners), (c) the resources needed to measure, (d) data quality and availability for measurement, (e) whether relevant evidence already exists, and (f) the timing of when the results are needed.

The measurement framework is also used to guide the identification of broad measurement approaches to addressing each prioritized question, which are a function of the level of rigor required to generate credible evidence on a given question. For example, questions that need a high degree of attribution, such as identifying *effective solution levers* that can be scaled up, are likely to need a considerably higher level of rigor in the measurement design than would a question related to effective implementation or understanding the scale-up process. This is because stakeholders such as the Government of Bihar will need convincing evidence to justify the (re)allocation of resources to the replication and scale up a new intervention delivery approach.

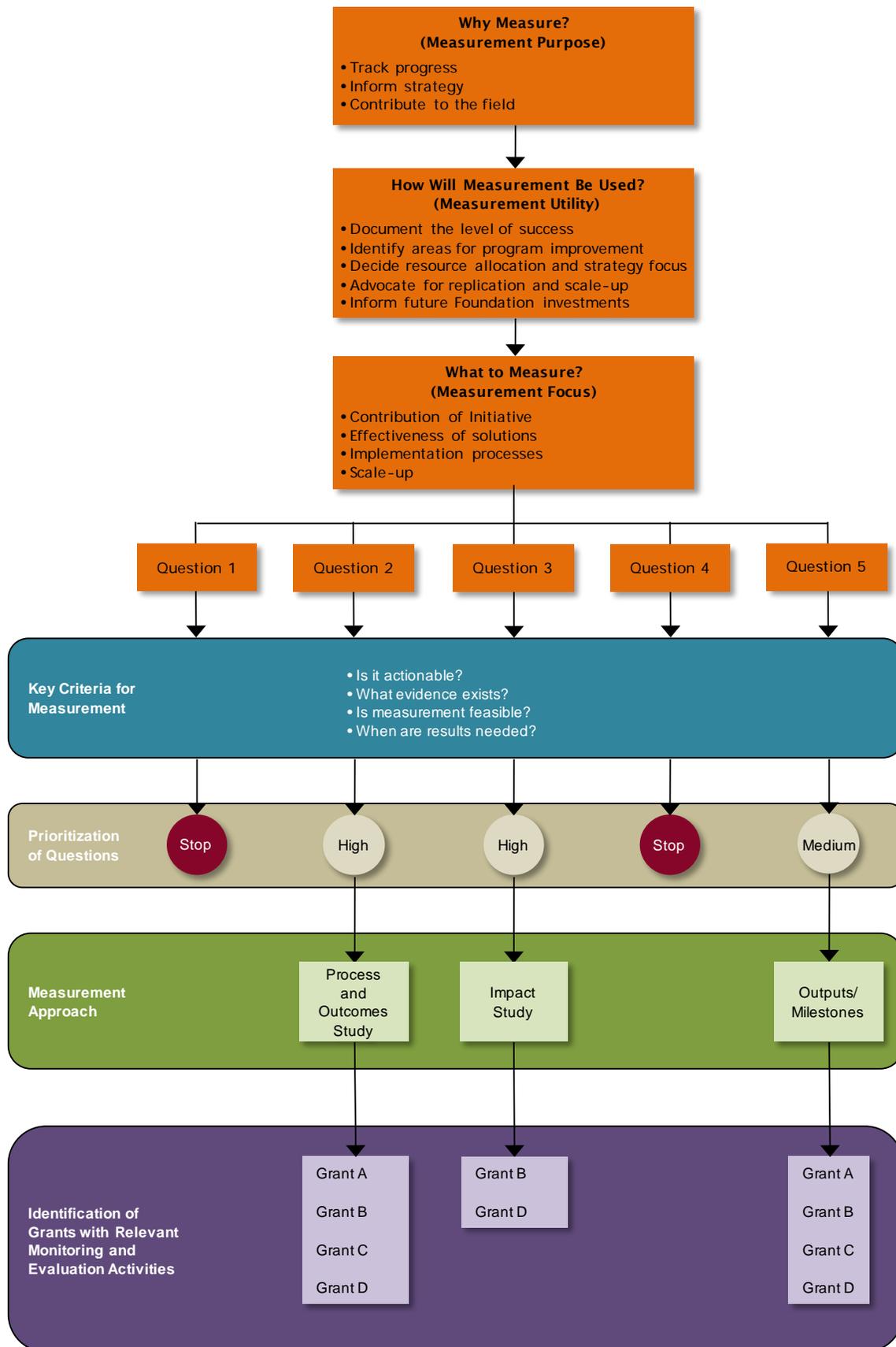
The measurement framework also maps the questions and measurement approaches to relevant program and/or MLE activities at the grant level. This mapping helps to identify the grant-level activities on which a question/evaluation may focus; whether grant-level monitoring and evaluation efforts (including data collection) may address a particular question, and the extent to which additional, external evaluation activities are needed.

2. Description of Measurement Framework Components

Next, we describe each of key component of the general measurement framework that we will use to guide the development of the Bihar Initiative’s MLE plan.

- **Why measure?** This step involves identifying the key objectives of measurement. Thinking about how these measurement objectives relate to the goals of a particular effort, such as the Bihar Initiative, will provide overarching context for how MLE results will be used and, therefore, provide guidance for the development of the MLE plan. We use the foundation’s *Guide to Actionable Measurement* to highlight three key measurement objectives:

Figure A.1. Measurement Framework



- **Track progress.** This measurement area emphasizes accountability and requires regular and systematic documentation of key aspects of a strategy’s progress toward meeting its goals and objectives, providing continual feedback on how the strategy is being implemented and its progress toward its targets. This type of measurement provides results that can be used for the strategy’s annual plan/update and mid-course correction.
 - **Inform strategy.** This measurement area emphasizes testing theories of change and key assumptions underlying a strategy, as well as understanding factors influencing success and failure. Although such information can be used in annual strategy review and planning, it is most suitable for making significant longer-term changes to the course of a strategy and its grants.
 - **Contribute to the field.** This measurement area emphasizes the dissemination of results to various stakeholders and foundation partners, and the generation of evidence that can be used to promote the replication and scale-up of effective solution levers.
- **How will measurement be used?** Closely tied to the question of “Why measure?” is how the information obtained will be used. Determining the primary use of MLE results will help identify measurement areas, select appropriate questions, and possibly indicate the evaluation approaches or methods to use. Potential uses of information include documenting success, identifying areas for program improvement, decision making related to resource allocation and focus, mobilizing support, and redistributing or expanding locations of the intervention.
 - **What to measure?** Using the processes described in the logic model for the Initiative, its underlying theory of change and key assumptions, and assessing the purposes of measurement and how the measurement will be used, we develop learning and evaluation questions. As noted in Section E, potential topic areas for measurement for the Bihar Initiative include:
 - Overall contribution of the Initiative to improving family health outcomes
 - Effectiveness and cost-effectiveness of innovative family health solutions
 - Implementation processes, successes, and challenges
 - Scale-up to the 8 priority districts and to the state level
 - **Key criteria for measurement.** When a set of learning and evaluation questions have been identified, each question can be assessed with respect to a set of criteria to determine its relevance for measurement. These criteria will aid in the prioritization of learning and evaluation questions, help determine the evaluation approach needed to answer these questions, and provide the desired level of certainty or attribution needed.
 - **Is it actionable?** MLE should address questions that will result in information that can be acted upon at the initiative or grantee level. It may also include questions that contribute to the field in order to promote the goals of generating convincing evidence for scale-up and replication.
 - **Is measurement feasible?** While measurement is often feasible, in some instances, a question might be of interest to the foundation and external stakeholders and provide new knowledge and learning, but it might not have

measurable outcomes or be feasible to answer in a credible manner within the available resources.¹²

- **Are high quality data available for measurement?** If high quality data cannot be collected at a reasonable cost, it will be important to assess whether the question will produce credible answers and whether alternative approaches or a slightly different question might need to be answered.
- **When are the results needed?** The time frame within which results are needed at the Initiative or grantee level will be a critical factor in determining whether certain outcomes can be prioritized for measurement. For example, long-term outcomes that require 5 to 10 years to observe might not be feasible for MLE if the time frame of the intervention is 5 years.
- **Prioritization of questions.** Based on the criteria for measurement, learning and evaluation questions can be ranked as high or medium priority. Ideally, low priority questions (or those not feasible to measure adequately) should not be part of MLE.
- **Measurement approach.** The next step is to identify possible measurement approaches to answer the prioritized learning and evaluation questions. Although we could potentially use more than one approach to answer a particular question, a question often lends itself to a certain type of approach. There are several broad types of evaluation approaches that we can use to answer the prioritized learning and evaluation questions, including randomized designs, comparison group designs, pre-post or trend analysis, implementation/process analysis and diffusion and network analysis, which can be utilized depending on the purpose and utility of measurement and what is feasible.
- **Using relevant grants.** After all relevant stakeholders agree upon questions and measurement approaches, we will identify the relevant grants that can provide information specific to an Initiative-level question (as well as whether the grants' monitoring and evaluation activity already contributes to the question or if additional measurement will be needed).

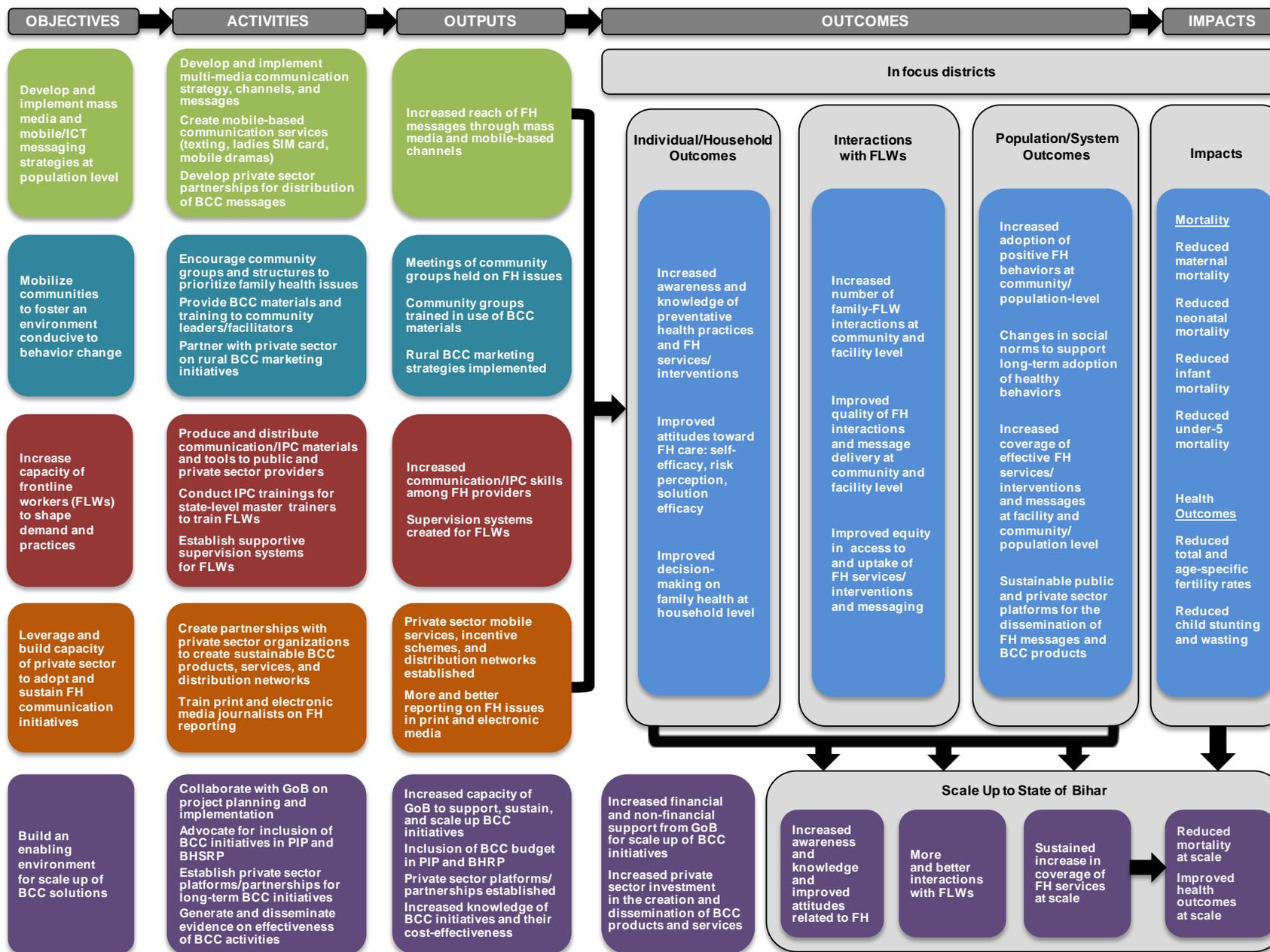
¹² Alternative approaches to informing the question should also be examined to consider whether some information is available for such questions. For example, a question might require high levels of attribution, but this might not be feasible given implementation plans (such as when an intervention is implemented district-wide or when the changes are at the system level). In such cases, alternative measurement, including assessing uptake or the process of adapting to systems changes, could still provide valuable information for the foundation and external stakeholders.

APPENDIX B

LOGIC MODELS FOR THE GRANTS

This page has been left blank for double-sided copying.

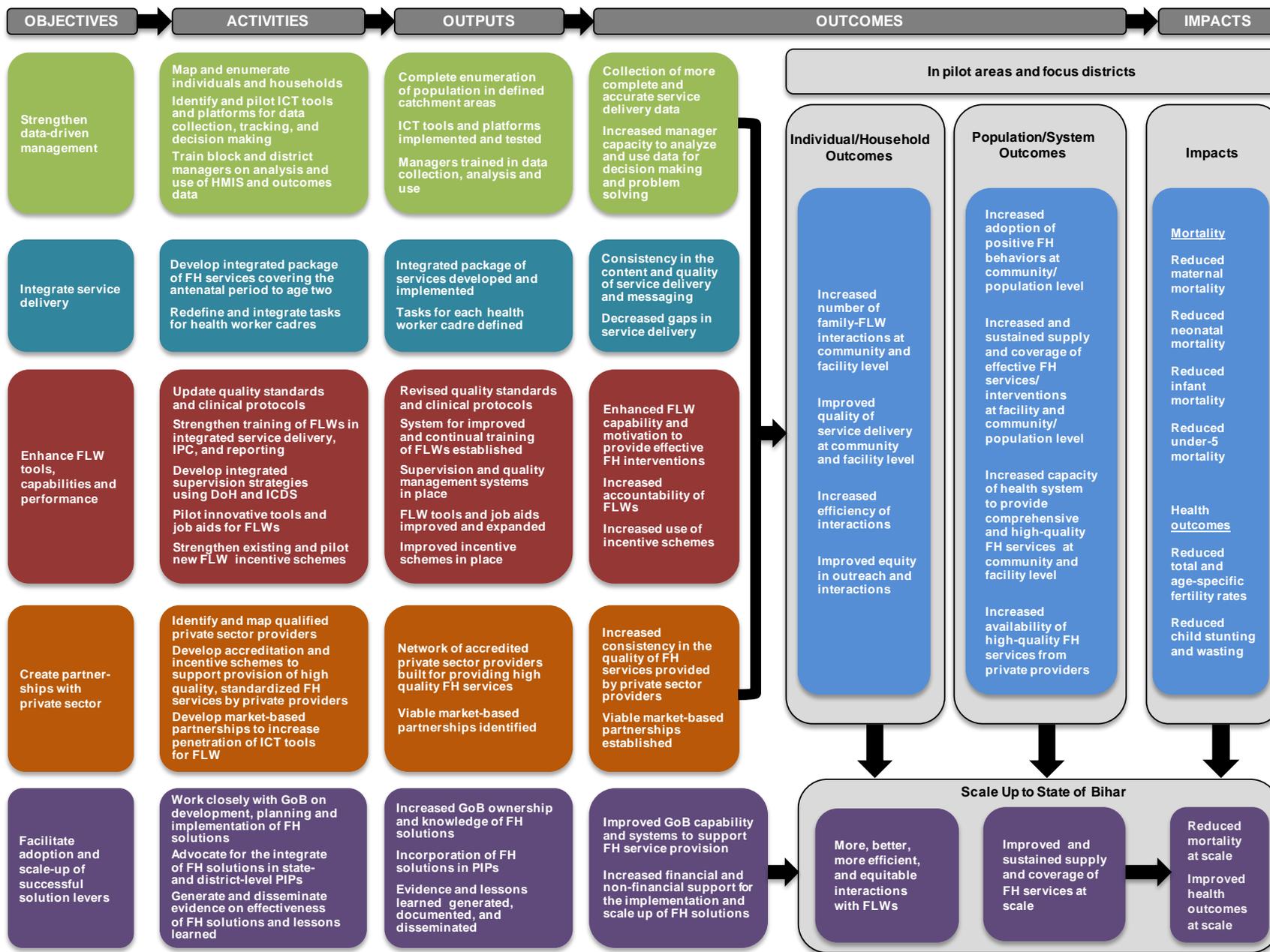
Figure B.1. Shaping Demand and Practices to Improve Family Health in Bihar (BBC World Service Trust)



This page has been left blank for double-sided copying.

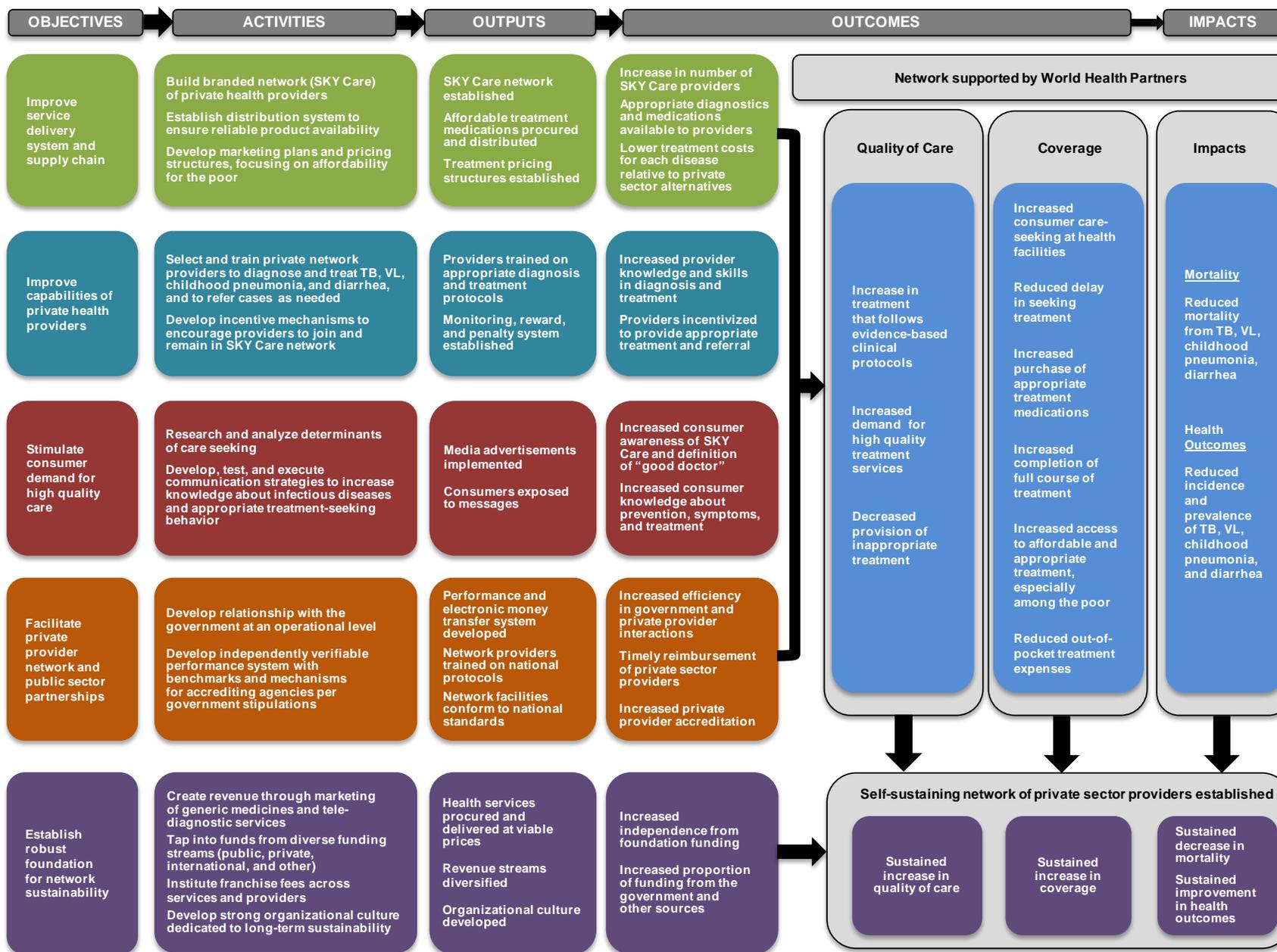
Figure B.2. Family Health Initiative in Bihar (CARE)

B.5



This page has been left blank for double-sided copying.

Figure B.3. Engaging Private Providers to Improve Management of Tuberculosis, Visceral Leshmaniasis, Childhood Pneumonia, and Diarrhea in Bihar (World Health Partners)



This page has been left blank for double-sided copying.

APPENDIX C

PRELIMINARY SET OF INDICATORS

This page has been left blank for double-sided copying.

Key Indicators for the Bihar Initiative

Measurement requires translation of the concepts or items from a logic model into specific *indicators* that can be used to measure progress toward the outputs and outcomes specified in the model, and to test theory of change and assumptions underlying a program's overall strategy. As part of the MLE plan, we will identify specific indicators that will be measured to address each of the evaluation questions included in the plan. Our first step in this process was to create a compilation of indicators from the foundation and grantee documents. In this Appendix, we describe our approach to compiling a preliminary list of relevant indicators for the Initiative, present these indicators, and discuss the process for identifying a final set of core indicators for the Bihar Initiative's MLE efforts.

1. Approach to Identifying Indicators

Using the original grant proposals and the draft Initiative-level log frame, we compiled a preliminary list of indicators that correspond to each output and outcome in the Initiative-level logic model (Table C.1). These indicators capture the grantees' and foundation's initial thinking on appropriate and key indicators, an important initial step in developing a comprehensive list of indicators from which priority indicators for measurement can be selected.

We had several reasons for starting out with this collation of indicators proposed by the foundation and grantees, and for creating an initial indicator repository of sorts. First, we wanted to build on the indicators that the foundation and grantees had already developed. Second, this process of mapping proposed indicators to components of the logic model provides us with a better understanding of what the foundation and grantees view as relevant and key indicators. Third, one of the requests we received from the foundation and grantees at the Bihar partners meeting in Delhi in December 2010 was to assess the extent to which there is overlap in proposed indicators across grantees, both in terms of specific indicators and the proposed targets for those indicators.

The Initiative-level indicator table (Table C.1) is organized by logic model component and contains indicators that are common across grantees, as well as indicators that are unique to a specific grant and to the overall Initiative.¹³ In the table, we include columns that indicate, for each specific indicator, whether the foundation or a particular grantee proposed measurement of that indicator for assessing progress towards or attainment of specific objectives.

For all indicators included in Table C.1, we indicate the data source proposed in the foundation or grant document for a specific indicator (in parentheses next to that indicator), when such information was available. If a specific source was not noted in any foundation or grantee documents, we did not include any proposed source. As we work on refining and prioritizing the indicators, we will also focus on potential data sources for measurement of each indicator, including existing data sources and primary data to be collected by the grantees. Current data sources proposed by the foundation and grantees fall into one of the following general categories:

¹³ We do not include output indicators in the Initiative level table, as outcomes will be the primary focus of measurement at this level. We might eventually include select output indicators in the Initiative level table, as part of efforts to track progress, including implementation of key activities.

- **Extant household survey data:** District Level Household Survey, Annual Health Survey, and National Family Health Survey.
- **Administrative data:** Management information systems, district plans and expenditures, internet and mobile technology use, television audience measurement, program monitoring and financial reports, training reports, community group reports and minutes, registered tuberculosis cases, the National Vector-Borne Disease Control Programme, the Revised National TB Control Programme, and sales and distribution data.
- **Survey data:** Grantee-initiated consumer surveys, facility surveys, patient surveys, household surveys, and market surveys.
- **Qualitative data:** Stakeholder interviews, focus groups, and review of publications.

It is important to note that the indicators included in Table C.1 do not represent a comprehensive or prioritized list of indicators to be measured as part of the MLE plan for the Bihar Initiative. The identification and prioritization of indicators to be included in and measured as part of the Bihar MLE plan will be driven by the key learning questions.¹⁴

2. Steps to Refine the Indicators

We will update the preliminary list of potential indicators presented here using information provided to us by the foundation and grantees during upcoming MLE discussions. After identifying a final set of learning and evaluation questions for the Initiative’s MLE effort, we will draw from and add to this evolving “indicator repository” to identify the most appropriate and critical indicators for addressing each of the MLE questions. We will share the resulting set of prioritized indicators for measurement with the foundation and grantees for feedback and, through an interactive and collaborative process, will identify a final set of core indicators for the Bihar Initiative’s MLE plan.

We will follow several guiding principles as we continue to identify and prioritize indicators for various measurement purposes:

- The indicator *has to be relevant in answering the prioritized learning and evaluation questions*. To the extent possible, we will rely on existing indicators that the grantees include in their documents, but it is possible that we will have to add additional indicators to answer specific learning questions.
- The indicator should be *relevant and important for the grantee* (either for program implementation or program improvement) and/or for the foundation.
- Where feasible, indicator should *use extant data*, which will also facilitate tracking after the project ends.
- The principles of *specificity, measurability, achievability, feasibility (realistic), and timeliness (SMART)* will drive additional data collection needs.

¹⁴ Further identification and prioritization of indicators will be driven by foundation’s and grantees’ needs for monitoring and tracking, data availability and quality, and learning and evaluation questions. We will also examine indicators and data sources for relevancy, reliability, accuracy, and redundancy.

Table C.1. Preliminary Indicators for Bihar Initiative

Logic Model Component	Indicator (Data Source)	BMG	CARE	BBC	WHP
INDIVIDUAL/HOUSEHOLD OUTCOMES					
Improved knowledge, awareness, and social norms related to family health	• Percentage of target groups with correct knowledge on at least three contraceptive methods	√			
	• Percentage of women with a birth preparedness plan (including at least 2 elements)		√		
	• Percentage of mothers-in-law with correct knowledge about antenatal care, Janani Suraksha Yojana, newborn and postpartum care	√			
	• Percentage of women with knowledge of two newborn danger signs (Household surveys)		√		
	• Percentage women of reproductive age and men with children aged 6-23 months who are aware about early, exclusive breast feeding and complementary feeding	√			
	• Percentage of women with children under 24 months old who are aware about the names and timing of vaccinations to be given to children by age 1	√			
	• Percentage of women and men with children under 24 months old who know the date and location for vaccination sessions	√			
	• Percentage of low income consumers that understand importance of complying with full course of treatment for infectious diseases (Consumer survey and focus groups)				√
	• Percentage of VHSC/community leaders who are knowledgeable about key antennal, newborn, immunization, birth spacing and infant feeding	√			
Increased quantity of family-FLW interactions at community and facility level	• Percentage of families receiving 6 or more healthcare contacts during pregnancy (AHS Woman Schedule; NFHS)		√	√	
	• Percentage of families receiving 6 or more healthcare contacts between 1 and 24 months of age		√	√	
	• Percentage of women and newborns receiving a postnatal evaluation at home within 48 hours after birth or coming back from facility (AHS Woman Schedule; DLHS; NFHS)		√	√	
Improved quality of interactions at community and facility level	• Percentage of women in the target groups who report FLWs as a major, trusted and influential source of information on family health (Household surveys)			√	
	• Percentage of families satisfied with specific <i>household, outreach</i> and facility-based services disaggregated by service, poverty-quintile, and caste		√		
	• Percentage of FLWs with correct knowledge about family health (FLW survey)	√			
	• % ANMs, ASHAs and AWWs who score 7 out of 10 point scale in terms of knowledge and communication skills related to MNH, nutrition, immunization and family planning	√			
	• Percentage of FLW who receive at least one supervisory contact per month (FLW survey)		√		
	• Percentage of AWCs with exact definition of catchment area and enumeration of families (FLW survey)		√		
	• Percentage of AWWs, ANMs and ASHAs who use new job aids		√		
	• Percentage of AWWs and ANMs who demonstrate use of BCC tools for conducting home visits	√			
	• Percentage of FLWs that adhere to defined quality standards for delivery of selected essential services—postnatal check-up mother/newborn, family planning counseling		√		
	• Percentage of deliveries with a skilled attendant with active management of the third stage of labor (AMTSL) meeting quality criteria (stratified by facility or home delivery)		√		

Table C.1 (continued)

Logic Model Component	Indicator (Data Source)	BMG	CARE	BBC	WHP
	<ul style="list-style-type: none"> Percentage of public and private facilities complying with quality standards during delivery and immediate postpartum care 		√		
Increased efficiency of interactions	<ul style="list-style-type: none"> Percentage of FLWs who captured >90% of pregnancies and births in focus districts 	√			
Improved equity in outreach and family-FLW interactions	<ul style="list-style-type: none"> Percentage reduction in disparity in receipt of selected services between lowest and highest poverty quintile and by caste: 2 birth planning lessons, skilled birth attendant, 2 postnatal visits in first week, full immunizations (AHS Woman Schedule; DLHS) 		√		
	<ul style="list-style-type: none"> Percentage of FLWs who prioritize contacting families from the lowest poverty quintile and scheduled castes 	√			
	<ul style="list-style-type: none"> Percentage of low income consumers in Bihar that recall key IPC messages (Consumer survey and focus groups) 				√
POPULATION/SYSTEMS OUTCOMES					
Increased adoption of positive FH behaviors at community/population level	<ul style="list-style-type: none"> Contraceptive prevalence rate 			√	
	<ul style="list-style-type: none"> Median age at birth of first child (AHS; DLHS NFHS) 	√			
	<ul style="list-style-type: none"> Birth interval between the first and second child (AHS, DLHS, NFHS) 	√			
	<ul style="list-style-type: none"> Percentage of women delivering with skilled birth attendant present initiate breastfeeding within 1 hour of delivery (AHS; DLHS; NFHS) 		√	√	
	<ul style="list-style-type: none"> Percentage of newborns with immediate skin-to-skin contact (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of infants exclusively breastfed for 6 months—including 70% exclusively breastfed except for water (AHS; DLHS; NFHS) 		√	√	
	<ul style="list-style-type: none"> Percentage of children 6-24 months old receiving appropriate complementary feeding—including continued breastfeeding, and age appropriate frequency quantity, and diversity of complimentary foods (AHS; DLHS; NFHS) 		√	√	
Increased and sustained supply of effective FH services/interventions and messages at facility and community level	<ul style="list-style-type: none"> Percentage of blocks with public or accredited private facilities that offer 7 basic emergency obstetric and neonatal care signal functions (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of public health centers with functional newborn corners (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of FLWs who reported stock-outs in the last six months on critical supplies—vaccines, supplementary food, vitamin A, iron, etc 	√			
	<ul style="list-style-type: none"> Percentage of target groups in Bihar exposed to at least one mobile/mass media intervention related to FH (Mobile usage data; Television Audience Measurement; Household surveys) 			√	
	<ul style="list-style-type: none"> Percentage of private providers carrying appropriately packaged dispersible amoxicillin for pneumonia and dispersible zinc for diarrhea at affordable price (Provider facility survey; Sales and distribution data; Patient survey) 				√
	<ul style="list-style-type: none"> Percentage of institutions that adhere to defined quality standards for delivery of selected essential services—e.g. active management of the third stage of labor, family planning services 		√		
	<ul style="list-style-type: none"> Percentage of 1st level facilities where glucocorticoids are available (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of 1st level facilities with functioning bag and mask (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of primary health center providers trained in asphyxia management (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of 1st level facilities with Helping Babies Breathe algorithm available (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of 1st level facilities with infection control guidelines established and implemented (Facility survey) 		√		

Table C.1 (continued)

Logic Model Component	Indicator (Data Source)	BMG	CARE	BBC	WHP
	<ul style="list-style-type: none"> Percentage of women with facility births with preterm labor given glucocorticoids (Facility survey) 		√		
	<ul style="list-style-type: none"> Percentage of facility births where a partograph was used correctly (Facility survey) 		√		
	<ul style="list-style-type: none"> Integration of private clinics and providers to expand the availability of services (Facility survey) 	√	√		
Increased coverage of effective FH services/ interventions at facility and community/population level	<ul style="list-style-type: none"> Percentage of women receiving of 3 or more antenatal check-ups during pregnancy (AHS Woman Schedule; DLHS; NFHS) 		√	√	
	<ul style="list-style-type: none"> Percentage of women delivering with a skilled birth attendant— facility plus home (AHS Woman Schedule; NFHS) 		√	√	
	<ul style="list-style-type: none"> Percentage of pregnant women with unmet needs for emergency obstetric care 		√	√	
	<ul style="list-style-type: none"> Percentage of infants delivered at facility that receive immediate needed newborn care— asphyxia management as needed, thermal care (AHS Woman Schedule; DLHS) 		√		
	<ul style="list-style-type: none"> Percentage of children fully immunized (AHS Woman Schedule; DLHS) 		√	√	
	<ul style="list-style-type: none"> Percentage of children under 24 months who received 2 doses of vitamin A and iron supplements 	√			
	<ul style="list-style-type: none"> Percentage of diarrhea episodes treated with ORS/zinc (DHLS; Patient survey) 				√
	<ul style="list-style-type: none"> Percentage of TB cases completing treatment per Revised National TB Control Programme (RNTCP) protocols (DLHS; RNTCP) 				√
	<ul style="list-style-type: none"> Percentage reduction in median treatment delay amongst confirmed TB cases (Survey of registered TB cases) 				√
	<ul style="list-style-type: none"> Percentage of confirmed TB cases completing approved treatment in network (WHP MIS) 				√
	<ul style="list-style-type: none"> Percentage of VL cases completing treatment with approved regimens (NVBDGP) 				√
	<ul style="list-style-type: none"> Percentage of children under 5 with symptoms of pneumonia treated with dispersible amoxicillin or referred as necessary (NFHS; Patient survey) 				√
Increased capacity of GoB and health system to provide integrated, comprehensive, and high-quality FH services at community and facility level	<ul style="list-style-type: none"> Percentage of state budget allocated to NRHM 	√			
	<ul style="list-style-type: none"> Percentage increase in NRHM allocations to districts for FH services and interventions (District plans and expenditures) 	√			
	<ul style="list-style-type: none"> Percentage reduction in human resource gaps in NRHM and ICDS 	√			
	<ul style="list-style-type: none"> Use of data-driven planning and review processes at facility, district and state level 	√			
	<ul style="list-style-type: none"> Percentage of districts engaged in quality management and quality improvement processes (District plans) 	√			
	<ul style="list-style-type: none"> Percentage of district PIPs that successfully incorporate FH service delivery solutions (District plans and expenditures) 	√			
	<ul style="list-style-type: none"> Percentage of District Project Planning Officers who are capacitated to develop effective behavior change communication solutions for inclusion in program implementation plans (Interviews) 	√			
	<ul style="list-style-type: none"> Percentage of all district PIPs that successfully incorporate communication solutions (District plans) 	√			
IMPACTS					
Reduced maternal mortality	<ul style="list-style-type: none"> Maternal mortality rate (AHS Mortality Schedule; NFHS) 	√			
	<ul style="list-style-type: none"> Case fatality rates post management of delivery-related complications (facility data) 	√			
Reduced neonatal mortality	<ul style="list-style-type: none"> Neonatal mortality rate (AHS Mortality Schedule; NFHS) 	√			

Table C.1 (continued)

Logic Model Component	Indicator (Data Source)	BMG	CARE	BBC	WHP
	<ul style="list-style-type: none"> Case fatality rates post management of complications in newborns (facility data) 	√			
Reduced infant mortality	<ul style="list-style-type: none"> Infant mortality rate (AHS Mortality Schedule; NFHS) 	√			
	<ul style="list-style-type: none"> Case fatality rates post management of complications in infants (facility data) 	√			
Reduced under-5 mortality	<ul style="list-style-type: none"> Under-5 mortality rate (AHS Mortality Schedule; NFHS) 	√			
Reduced fertility rates	<ul style="list-style-type: none"> Total fertility rate (AHS Woman Schedule; NFHS) 	√			
	<ul style="list-style-type: none"> Age specific fertility rates among women aged 15-30 (AHS Woman Schedule; NFHS) 	√			
Reduced child stunting and wasting	<ul style="list-style-type: none"> Percentage of stunting among children aged 24-36 months (NFHS) 	√			
	<ul style="list-style-type: none"> Percentage of wasting among children aged 6-36 months (NFHS) 	√			
Reduced morbidity and mortality related to TB, VL, childhood pneumonia, and diarrhea	<ul style="list-style-type: none"> Incidence and prevalence of TB, VL, childhood pneumonia, and diarrhea 				√
	<ul style="list-style-type: none"> Case fatality rates of TB, VL, childhood pneumonia, and diarrhea 				√

- AHS = Annual Household Survey
- ANM = Auxiliary Nurse Midwife
- ASMA = Accredited Social Health Activist
- AWW = Anganwadi Worker
- DLHS = District Level Household Survey
- FLW = Frontline Workers
- ICDS =Integrated Child Development Services
- NFHS = National Family Health Survey
- NRHM = National Rural Health Mission
- NVBDCP = National Vector Borne Disease Control Programme
- PIP = Program Improvement Plan
- RNTCP = Revised National TB Control Programme
- VHSC = Village Health and Sanitation Committee

This page has been left blank for double-sided copying.

MATHEMATICA
Policy Research, Inc.

www.mathematica-mpr.com

Improving public well-being by conducting high-quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research