Measurement, Learning, and Evaluation for the Ananya Program (Family Health Initiative in Bihar)

Design Report

October 31, 2011

Kimberly Smith
Anu Rangarajan
Evan Borkum
Lalit Dandona

MATHEMATICA
Policy Research

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

Contract Number:
4993/18924(13)

Mathematica Reference Number:
06893.500

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 for the Ananya Program (Family Health Initiative in Bihar)

Design Report

October 31, 2011

Kimberly Smith
Anu Rangarajan
Evan Borkum
Lalit Dandona

MATHEMATICA
Policy Research

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

CONTENTS

TABLES

FIGURES

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

EXECUTIVE SUMMARY

Program Overview

In 2010, the Bill & Melinda Gates Foundation launched the Family Health Initiative in Bihar, India (now named "Ananya", a Sanskrit word meaning "unique" or "unlike others"). The goals of the Ananya program (2010–2015) are to reduce maternal, newborn, and child mortality; malnutrition; fertility; and morbidity from infectious diseases by developing and implementing innovative and integrated health solutions that involve both the public and private sectors. More specifically, the program 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 its goals, the foundation has adopted a broad-based family health approach that addresses both supply- and demand-side barriers to increased uptake, coverage, and quality of family health interventions through a synergistic set of four complementary grants. Two of these grants aim to strengthen the supply and delivery of essential family health and infectious disease interventions in the public and private sectors, respectively. One demand-side grant focuses on changing behaviors, social norms, and self-efficacy to support family health through a multichannel communication strategy. The fourth, a community mobilization grant, focuses on strengthening and creating community structures to support the delivery and uptake of family health services.

A key component of the Ananya program is the scaling up of successful approaches to the delivery of effective family health and infectious disease interventions. To this end, the Ananya program will focus initially on implementing and assessing the effectiveness of an integrated package of demand- and supply-side approaches to improving family health outcomes in select focus districts (8 districts for all grants, except for the private sector grant, which will focus initially on 12 districts). Based on the findings and lessons learned from these efforts, the program will then promote and facilitate the replication and scale-up of successful strategies by the GoB and other development partners in the remaining 30 (of 38) districts in Bihar.

To achieve its scale-up and impact goals, the Ananya program focuses on leveraging and strengthening existing public and private sector delivery platforms, including state and local government health initiatives, such as the National Rural Health Mission (NRHM) and the Bihar Health Sector Reform Project (HSRP) funded by DFID. As part of this effort, the foundation has signed a memorandum of cooperation with GoB and established several forums to facilitate coordination and collaboration across key development partners.

MLE Overview and Questions

The foundation has contracted with Mathematica Policy Research to lead the measurement, learning, and evaluation (MLE) component of the Ananya program. Mathematica will work closely with its lead India evaluation partner, the Public Health Foundation of India (PHFI), in designing and executing the MLE plan. Mathematica is also partnering with Sambodhi, an Indian organization, which will lead the primary data collection effort and participate in other evaluation activities. The foundation has made awards to two other MLE partners—COHESIVE-India and the Center for Global Health Research (CGHR)—with whom Mathematica will coordinate. COHESIVE-India will focus primarily on the evaluation of the grant that focuses on private sector provision of infectious

disease services; CGHR will focus on compiling district-level estimates from existing secondary data to answer select learning and evaluation questions.

The Ananya MLE effort will focus on addressing three broad categories of research questions:

1. What family health approaches were implemented under Ananya? Did they achieve scale?
2. Did the implementation of these approaches at scale have impact? Were the highly innovative approaches tested by grantees effective?
3. What was the cost of implementing the program and were these costs effective?

In answering these questions, the MLE component of the Ananya program aims to (1) inform the foundation, Government of Bihar (GoB), and other partners about implementation progress, successes, and failures; (2) provide information to guide mid-course corrections; and (3) assess the impact of the overall program and the innovative approaches and solutions being implemented under it. Evidence generated through the MLE effort will provide critical information to the foundation, grantees, GoB, and external stakeholders to guide decision making related to program improvement and whether and how to scale-up innovative family health approaches in order to achieve the Millenimum Development Goals 4, 5, and parts of 6s.

MLE Approach

The MLE plan consists of three main evaluation components, each of which will inform answers to MLE questions above: (1) process and scale-up analysis, (2) the impact analysis, and (3) cost and cost-effectiveness analysis. These components are very closely integrated.

1. Process and Scale-Up Analysis

The MLE effort will include a rigorous implementation or process evaluation, which will provide a comprehensive picture of what interventions are being implemented, how they are being implemented, and the factors affecting implementation. The purpose of this MLE component is to generate information to (1) understand implementation processes, successes, and failures that will inform program improvement; (2) inform our interpretation of impact analysis findings; and (3) guide replication of program innovations by GoB and other development partners. Given the importance to Ananya's success of achieving scale, a key component of the process evaluation component will be to measure the extent to which scale-up occurred, understand and document the scale-up process, and identify the factors that facilitate and inhibit scale-up.

The timing of the process and scale-up analyses will be aligned with program implementation on the ground. Three rounds of process evaluation are anticipated: fall 2012, fall 2013, and fall 2015. Each round will involve field visits to select districts, blocks and villages, as well as the analysis of secondary data—such as project MIS, HMIS, administrative data, and district plans. The first round of process evaluation will occur roughly 12 to 15 months after implementation start-up for most grantees and will focus on characteristics of the interventions being implemented and how they are being rolled out; the context in which they are being rolled out; the acceptability, uptake, and perceived benefits of the interventions among beneficiaries and other stakeholders; and any initial bottlenecks or challenges. The subsequent rounds will additionally focus on implementation progress, including achievement of key outputs and outcomes and the extent to which scale-up is occurring.

2. Measuring the Impacts of the Ananya Program and of Select Innovative Solutions

The ultimate goal of the Ananya program is to improve family health outcomes in Bihar. Therefore, measuring changes in key outcomes over the course of the program life cycle, and assessing the contribution of the program to these changes, is a key component of the MLE plan. However, designing a rigorous impact evaluation that can rigorously "attribute" changes in family health outcomes to the Ananya program—or to particular program components—is challenging. Bihar's health sector is a multiplayer environment in which there are several parallel health sector initiatives and programs being implemented or planned by the GoB and various donors. This context makes it difficult to attribute changes in outcomes to the Ananya program alone. In addition, the program itself consists of a variety of coordinated and synergistic interventions implemented simultaneously in the same set of districts in Bihar and targeting many of the same behaviors. This makes it challenging to isolate the impact of particular program components.

Therefore, the evaluation will focus on measuring the effects of the integrated package of approaches being implemented by grantees and their partners, as well as the extent to which the program achieved its overall goals in terms of changes in key outcomes and impact indicators. Although our attempts to rigorously attribute changes to the Ananya program itself will be limited, we will attempt to obtain some degree of plausible attribution using a comparison group design to assess the short-term effects of the program in the 8 focus districts. We will also attempt to disentangle the impact of specific program components to the extent possible, by examining changes in proximal outcomes closely related to specific program components and by relying on the qualitative process study component of the MLE effort. In addition, we plan to rigorously test the effectiveness of a small number of (between one to three) specific high-risk, high-reward innovations that will be implemented on a pilot basis in select areas, on top of the core package of demand- and supply-side interventions.

Impacts of the Ananya Program. The success of the Ananya program will be assessed at two key junctures in the program's lifecycle: (1) after scale-up has occurred in the 8 focus districts (2012) and (2) at the end of the program (2015), when statewide scale-up will have occurred and the overall effects of the program can be assessed.

A comparison group design will be used to assess the short-term effects of the program in the eight focus districts. One of the challenges in selecting comparison districts early on in the process is that, in the dynamic context of Bihar, a selected comparison district may end up getting health services intervention through a new development partner over the next year or two. To avoid selecting comparison districts that may no longer serve as a valid comparison group, we propose to select comparison districts—based on baseline measures—just prior to the time of the midline assessment. Although we will still use baseline data to select matches, selecting the comparison districts in 2013 prior to when the mid-line survey is conducted, will allow us to take into account any changes in the scope and intensity of activities by other development partners in the various districts of Bihar. Again, although the comparison group design will not allow us to attribute any changes to the Ananya program with certainty, it will improve the degree of attribution relative to a simple pre-post design, provided that a reasonable set of comparison districts can be identified using rigorous methods, and that extensive sensitivity testing is done with the results.

Given plans to scale up successful approaches statewide, a comparison group design will not be feasible for the overall assessment of the program's success at the end of the current five-year program cycle. In addition, because successful scale-up depends on the replication of effective family health approaches by other development partners, attribution of changes in outcomes to the

program alone are also not necessary. Instead, the second-stage impact analysis will assess whether Ananya met its objectives in terms of targeted changes in key indicators using a pre-post design to measure changes in outcomes between 2011 and 2015.

Due to limitations in existing data sources, primary data will be collected baseline, midline, and endline data for the impact analysis. The broad scope of the Ananya program—which is designed to improve a range of demand-side, supply-side, and health outcomes across multiple family health domains—necessitates a similarly broad-based data collection effort. The primary data sources for the impact analysis will include household, frontline worker, and facility/provider surveys. Data will be collected on a range of outcomes, including individual-level knowledge, attitudes and practices; family-frontline worker and provide interactions; and measures of population coverage of key family health interventions. Measurement of the impact of Ananya on mortality will focus on neonatal mortality, since grantees' interventions focus most on the neonatal period.

Effectiveness of highly-innovative solutions. The MLE effort will also include rigorous pilot testing of one to three highly-innovative solutions that are expected to add value to the program's core interventions. At the time that this report was written, the grantees were working with the foundation to identify a set of value-added solutions to be piloted and potentially rigorously-tested in the eight focus districts. Innovative solutions that have been recommended for rigorous testing include the use of mobile technology to improve the provision of care by frontline workers, using team-based incentives at the sub-center level to improve the quality of services provided by frontline workers, and the introduction and promotion of micronutrient power for children and postnatal mothers. The foundation will make a decision by November 2011 on which of these proposed solutions (or other ones) should be considered for rigorous testing, after which we will work with the grantees to design and conduct a rigorous evaluations of the selected interventions.

3. Cost and Cost-Effectiveness Analysis

Although evidence of effectiveness is critical for persuading development partners to adopt a new public health approach, the costs associated with replicating a program or approach and its impacts are also a key factor in replication and scale-up decisions. The costing component of the MLE plan will aim to generate estimates of overall program costs, the costs of major program components, the key cost drivers, and replication costs. In addition, it will attempt to determine the cost-effectiveness of the program and how cost-effectiveness evolved over the course of the program.

The costing study will be conducted from the perspective of the foundation (or program funder, and will employ, to the extent possible, an activity-based costing methodology and an "ingredients" approach to cost estimation. It will draw on both retrospective and prospective data from a variety of sources for the costing study, including financial records, program MIS data, key informant interviews, and administrative data. Some aspects of data collection for the costing study will be embedded in the impact and process evaluation data collection.

Dissemination of Results

The MLE component of the Ananya program will produce a wealth of information about the implementation, cost, and effects of the program over the next five years. The study's many and diverse findings will be communicated in an effective and timely manner to various stakeholders, including the GoB, foundation, grantees, community and development partners in Bihar, donors and NGOs working in the family health arena, researchers, and other members of the international

health community. We anticipate producing a range of dissemination products, including reports, topic papers, research/evaluation briefs, journal articles, and targeted emails; we will also present findings at various meetings and conferences.

Over the next few months, working closely with the foundation, we will develop a communications and dissemination plan for the project life cycle that reaches and engages key stakeholders, disseminates results in real time, and promotes feedback from and interaction with internal and external audiences. The plan will identify key internal and external audiences for the various MLE results and products. It will also create a tiered priority structure to categorize the various audiences, and the types of information or levels of (technical) detail desired by the targeted audience to ensure that the dissemination approach addresses the needs all key stakeholders.

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

I. INTRODUCTION

The belief that everyone should have the "opportunity to live a healthy productive life" guides the work of the Bill & Melinda Gates Foundation (the foundation). As part of its mission, the foundation's Global Health Program is committed to reducing illness and mortality in developing countries through the sustainable delivery of life-saving tools, technologies, and approaches. Most of the foundation's global health investments fall into two areas—family health and infectious diseases—that disproportionately affect developing countries. In the family health area, the foundation promotes innovative and integrated solutions for family planning; nutrition; maternal, neonatal, and child health; and vaccine-preventable diseases. The foundation's infectious disease work focuses on developing ways to prevent, manage, and treat diarrheal diseases, HIV/AIDS, malaria, pneumonia, tuberculosis, and neglected and other infectious diseases.

The Family Health Initiative in Bihar, India (now named "Ananya", a Sanskrit word meaning "unique" or "unlike others") 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 (MDGs) 4, 5, and elements of 6. In particular, Ananya takes an integrated demand- and supply-side approach to improving reproductive, maternal, neonatal, and child health that leverages resources and lessons learned from several of the foundation's Global Health program 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; Neglected and Other Infectious Diseases. The Water, Sanitation, and Hygiene strategy from the Global Development program has also joined this initiative, with a particular focus on affecting behavior change in the hand washing and hygiene areas as part of the demand side efforts under the initiative.

The foundation has contracted with Mathematica Policy Research to lead the measurement, learning, and evaluation (MLE) component of the Ananya program. Mathematica will work closely with its lead India evaluation partner, the Public Health Foundation of India (PHFI), in designing and executing the MLE plan. Mathematica will also partner with Sambodhi, an Indian organization, which will lead the primary data collection effort and participate in other evaluation activities. The foundation has made awards to two other MLE partners—COHESIVE-India and the Center for Global Health Research (CGHR)—with whom Mathematica will coordinate. COHESIVE-India will focus primarily on the evaluation of one grant that focuses on private sector provision of infectious disease services; CGHR will focus on compiling district-level estimates from existing secondary data to answer select learning and evaluation questions.

As part of the development of the MLE plan, Mathematica produced an "MLE Framework" report in April 2011, which included results frameworks (and logic models) for the overall program and the individual grants awarded under it, preliminary learning and evaluation questions to be addressed as part of the evaluation, and broad approaches to answering these questions (Rangarajan et al. 2011). Based on feedback and input from the foundation and grantees on the framework report and subsequent discussions with various Ananya partners, we developed a prioritized set of learning and evaluation questions on which the MLE effort will focus:

1. What family health approaches were implemented under Ananya? Did they achieve scale?
2. Did the implementation of these approaches at scale have impact? Were the highly innovative approaches implemented by grantees effective?
3. What was the cost of implementing the program and were these costs effective?

In answering these questions, the MLE component of the Ananya program aims to (1) inform the foundation, Government of Bihar (GoB), and other partners about implementation progress, successes, and failures; (2) provide information to guide mid-course corrections; and (3) assess the impact of the overall program and select innovative approaches and solutions being implemented as part of Ananya. Evidence generated through the MLE effort will provide critical information to the foundation, grantees, GoB, and external stakeholders to guide decision making related to program improvement and whether and how to scale-up innovative family health approaches in order to achieve the MDGs.

This report presents our approach to addressing the above three sets of broad learning and evaluation questions. In it, we present more detail on these questions and describe how we plan to use a range of primary and secondary data to provide comprehensive answers to them. Our goal is to provide sufficient detail so our overall MLE plans are clear. However, the design for some of the study components of the MLE plan (particularly the process, scale up, and costing components) will be further developed and refined as the grantees finalize their implementation plans, and prior to when we begin executing specific MLE activities.

Before presenting our MLE plan for the Ananya program, in the remainder of this chapter, we provide an overview of the Bihar context, the Ananya program and its grant portfolio, and key challenges for the MLE effort. We conclude this chapter with a road map for the rest of the report.

A. 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 (USD $1,070).[1] 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 underweight children.[2] 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.

[1] 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.

[2] Bill & Melinda Gates Foundation Request for Letter of Inquiry, Family Health Initiative for Bihar (2010–2015).

Under strong government leadership, 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 sectors, including physical infrastructure, education, and health. Despite these strides, the health status of the Bihari population, particularly those residing in rural areas, still requires 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 Rs. 600 crore (USD $135 million) to reduce maternal and child deaths, undernutrition, and unwanted pregnancies/fertility through increased scale and functionality of health services, systems strengthening, and greater engagement of nongovernment actors. In addition, the United Nations Children's Fund (UNICEF) has been operating in Bihar for several years to scale up comprehensive newborn care, strengthen routine immunizations, and address human resource shortages.

B. Description of the Ananya Program and Its Grant Portfolio

The overall goal of the foundation's Ananya program (2010–2015) is to reduce maternal, newborn, and child mortality; malnutrition; and fertility rates by developing and implementing innovative and integrated health solutions that involve both the public and private sectors. More specifically, the program 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.

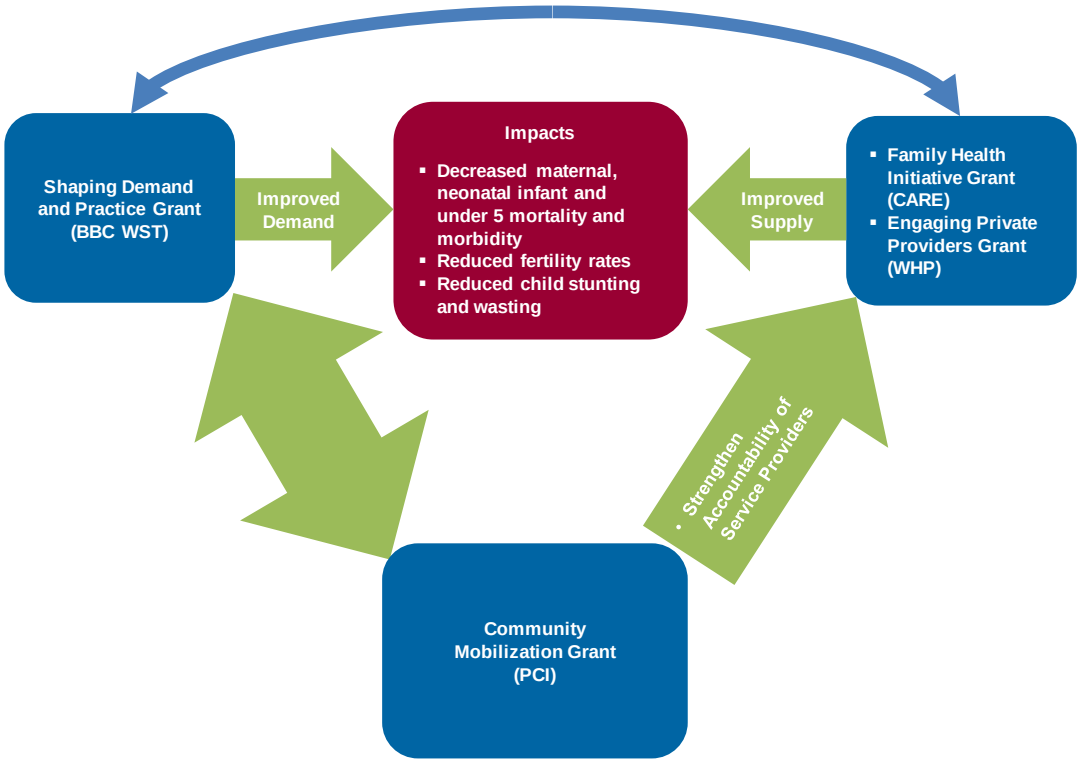
To achieve its goals and objectives, the foundation has adopted a broad-based family health approach that addresses both supply- and demand-side barriers to increased uptake, coverage, and quality of family health interventions and services through a synergistic set of four complementary grants. Two of these grants aim to strengthen the supply and delivery of essential family health and infectious disease interventions in the public and private sectors, respectively. One demand-side grant focuses on changing behaviors, social norms, and self-efficacy to support family health through a multichannel communication strategy. The fourth, a community mobilization grant, focuses on strengthening and creating community structures to support the delivery and uptake of family health services. Three of the four grants were awarded in fall 2010; the fourth (community mobilization) grant was recently awarded in October 2011:

1. ***Integrated Family Health Initiative (IFHI) in Bihar.*** IFHI, led by CARE, will work closely with the GoB to scale up innovative supply-side approaches to improving the coverage, quality, and uptake of critical family health services in Bihar. The approaches can be broadly categorized as strengthening data-driven management; integrating the delivery of family health services; improving the capabilities of and tools for frontline workers; and creating partnerships with private sector family health providers to extend the reach and quality of care. During the first two years of the initiative (2011 and 2012), IFHI will focus on developing and implementing a core set of supply-side interventions in eight focus districts. In addition, two or three highly innovative, value-added approaches (to be implemented on top of the core set of interventions) will be selected for rigorous pilot testing in select areas of the eight focus districts. In years 3 through 5 of the program, IFHI will support the GoB to scale up successful family health approaches implemented under the project to the entire state of Bihar.

2. ***Shaping demand and practices to improve family health in Bihar.*** The Shaping Demand and Practices (SDP) grant, being implemented by the BBC World Service Trust (BBC WST), aims to increase demand for key family health services and improve family health practices by increasing knowledge, changing attitudes, shaping social norms, and improving self-efficacy. It focuses on developing and testing innovative ways of providing information on key family health topics through various communication channels, including mass media, mobile and internet technology, community groups, and family–frontline workers interactions. The project aims to leverage and partner with the commercial private sector, as well as build the capacity of GoB, to sustain and scale up successful behavior-change communication approaches to improve family 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, being implemented by World Health Partners (WHP), aims to improve the availability and 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 in children under 5. Specific strategies for achieving these objectives include strengthening 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 ensuring the sustainability of the private provider network.
4. ***Community mobilization.*** The community mobilization grant was recently awarded to Project Concern International (PCI). By catalyzing collective community action to promote shifts in social norms and behavior change, PCI aims to improve family health and enhance the accountability and equity of health services across Bihar. PCI's program model will be implemented in three stages. The first stage will focus on creating and strengthening community organizations and networks in an effort to promote community accountability of health, sanitation, and welfare services. In the second stage, PCI will facilitate community audits of service delivery; promote the participation of community leaders in village-level governance and oversight structures; increase community awareness of equity, quality, and availability gaps; and build community linkages with media and local power structures. Finally, PCI will analyze and test models for state-wide scale-up of successful community mobilization intervention, while continuing to build partnerships to promote the sustainability of these interventions.

Although each of these grants has a different primary focus, they are intended to be complementary and, through coordination and synergies across grants, the foundation aims to increase the coverage of critical and efficacious interventions and ultimately reduce maternal and child mortality and morbidity. Figure I.1 provides a visual representation of the interaction among the grants and how they are expected to affect health outcomes in Bihar.

Figure I.1. Integrated Set of Demand- and Supply-Side Activities to Achieve Impacts



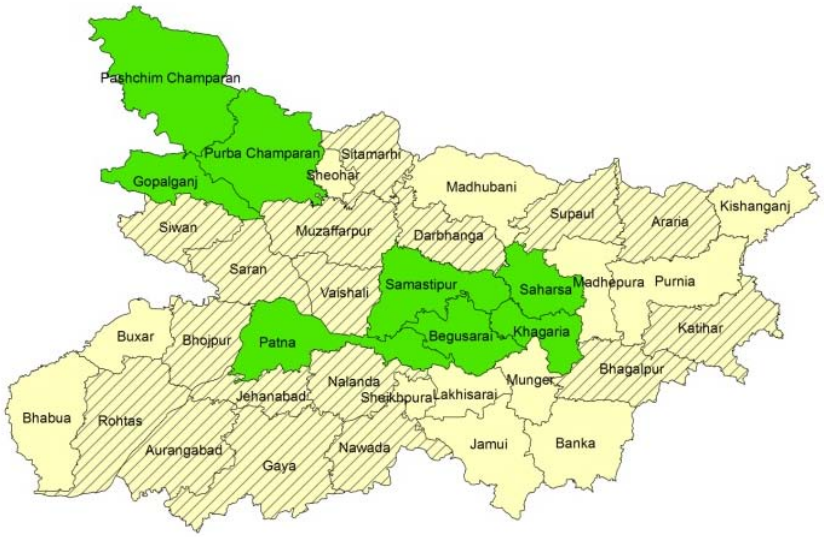
A key component of Ananya is the scaling up of successful approaches to the delivery of effective family health and infectious disease interventions. To this end, the Ananya program will focus initially on implementing an integrated package of demand- and supply-side approaches to improving family health outcomes in eight focus districts (Figure I.2). Based on the lessons learned from the implementation of these integrated interventions, the program will promote and facilitate the replication and scale-up of successful strategies by the GoB and other development partners in the remaining 30 (of 38) districts in Bihar.[3]

To achieve its scale-up and impact goals,[4] the Ananya program focuses on leveraging and strengthening existing public and private sector delivery platforms, including state and local government health initiatives, such as the National Rural Health Mission (NRHM) and the Bihar Health Sector Reform Project (HSRP) funded by DFID. As part of this effort, the foundation has signed a memorandum of cooperation with GoB that formalizes 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. In addition, the foundation has established several forums to facilitate coordination and collaboration across key development partners during the planning, implementation, and scale-up phases of the Ananya program.

[3] The WHP grant plans to implement in 12 districts in the first year (2011) and to scale up to 25 districts in the second year (2012).

[4] An operational definition of scale in the Ananya context will be developed as part of the process evaluation component of the MLE effort, with the input of the grantees.

Figure I.2. Geographical Coverage and Scale-Up of the Bihar Initiative



Districts of Bihar

Eight focus districts, scale-up targeted by 2012

Remaining districts, scale-up targeted by 2015

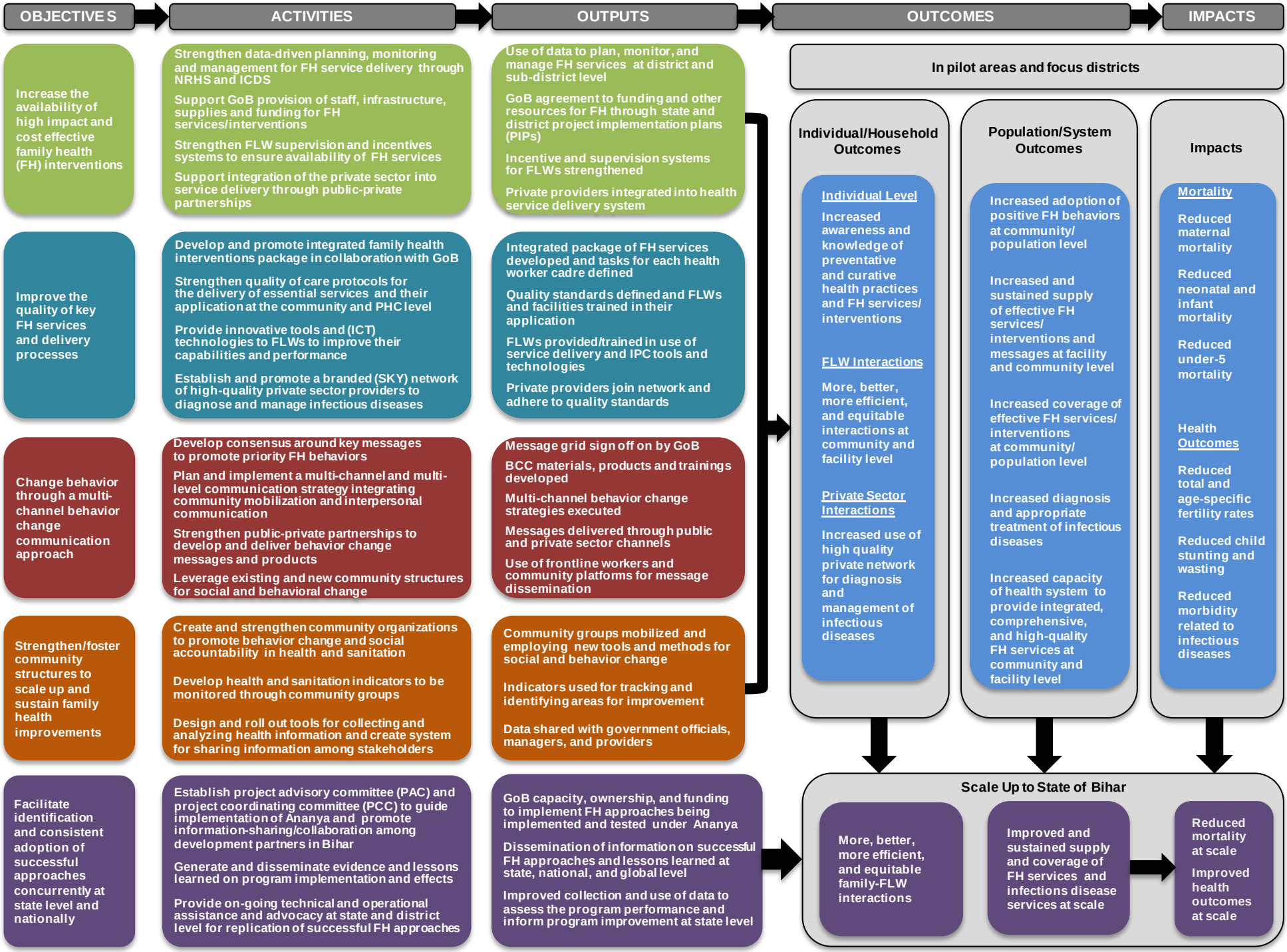
Note: BBC-WST and CARE grants will scale up throughout the state of Bihar by 2015. WHP plans to scale up to 25 districts by 2012; these include the 8 focus districts and the 17 districts marked with diagonal lines.

C. Logic Model of the Ananya Program

Figure I.3 presents the logic model for the overall Ananya program.[5] 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 linkages between each component of the logic model will be examined as part of the MLE plan to assess the validity of the program's theory of change and identify the contribution of program activities to outcomes. By identifying these linkages, as well as the key activities and outcomes for assessment, the logic model provides critical input for the MLE plan.

[5] Logic models for the three grants awarded in the last quarter of 2010 are included in Appendix A.

Figure I.3. Logic Model for the Ananya Program



- **Objectives.** The logic model includes five interrelated and synergistic program-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 through a multi-channel behavior change communication approach; (4) to strengthen community structures 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.[6]

- **Activities.** Activities reflect the inputs being implemented by the grantees as part of the Initiative. An illustration of an activity to increase the availability family health interventions is strengthening data-driven planning, by conducting full enumeration of frontline worker catchment areas and scheduling and planning for family health service delivery. Similarly, an activity to improve the quality of family health services is to strengthen quality of care protocols and their application at the primary health center (PHC) level, and to provide continual training to frontline workers to increase the number and quality of interactions between families and frontline workers.

- **Outputs.** The activities implemented by the grantees, in collaboration with the GoB and other partners are expected to result in outputs that indicate the direct and tangible execution of the activities. Outputs are the *direct and immediate result* of program activities and can be described as meetings held, plans developed, materials produced, trainings conducted, capacity built, policies implemented, and research conducted by the program. Examples of outputs include the development of behavior change communication materials and tools and the training of frontline worker in the use of information communication technology tools. Because the specific activities implemented by grantees vary by objective, outputs also differ across objective areas. Measurement of outputs is used to assess the program's implementation progress.

- **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, first within the selected focus districts, and following state-wide scale up, at the state level. Outcomes include changes in attitudes, knowledge, behaviors, skills, family-frontline worker interactions, and care processes. For the evaluation, measurement of these program outcomes contributes to the assessment of program effectiveness and ultimate impact.

  In the logic model, outcomes are presented in two categories:

  1. *Individual/household and frontline worker outcomes.* These include outcomes related to individuals' knowledge, attitudes, and behaviors related to family health and to changes in families' interactions with frontline workers, including the number, quality, efficiency, and equity of interactions. They also include families' use of and interaction with SKY network providers for the diagnosis and management of infectious diseases.

[6] Note that throughout the logic model, as well as the report, we define "family health" very broadly to include a range of family health domains, including maternal and reproductive health, nutrition, family planning, immunizations, infectious diseases, and sanitation and hygiene.

2. *Population/systems outcomes.* These outcomes demonstrate the reach of the program beyond the individual and interactions level to the facility, community, population, and systems level. They are related to changes in the adoption of key 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. Through sustained implementation and coverage of family health innovations at scale, leading to more and better interactions with frontline workers and at public and private facilities at the population level, the Ananya program aims to reduce maternal, neonatal and child mortality, fertility child malnutrition, and morbidity related to infectious diseases.

While the logic model does not explicitly capture the dynamic nature of the Ananya program, the various inputs and actors, and interrelationships between each of the objectives, activities, outcomes, and outcomes, it provides a useful overview of key program activities and measurement areas, and guides the development of learning and evaluation questions for the MLE effort.

D. Key Challenges for the MLE Effort

The approach to implementing program services by the grantees and the context in which it is being implemented largely influence what can and cannot be addressed by the MLE of the Ananya program. Our ability to answer the prioritized learning and evaluation questions for the MLE effort will be influenced by several contextual, environmental, and design challenges. We discuss some of the more important challenges that we expect to face in implementing the MLE of the Ananya program, and discuss our approaches to addressing them in the MLE plan,and provide an overall sense of what we expect the evaluation can address and what it cannot address.

- ***Complex and dynamic health sector landscape in Bihar makes it difficult to design an evaluation that provides "rigorous" attribution in observed changes to Ananya alone.*** In addition to the Ananya program, there are several parallel health sector initiatives and programs being implemented or planned in Bihar that aim to improve family health outcomes. The GoB is an active player in the health sector in Bihar and has invested in a variety of health sector reforms and initiatives across the state. In addition, donors such as DFID and UNICEF are supporting major health sector initiatives, with a focus on select districts. This multiplayer environment makes it challenging to design a rigorous impact evaluation that can attribute changes in family health outcomes to the Ananya program alone. While a rigorous evaluation with attribution will not be feasible, we will use quantitative and qualitative methods to understand the effects of Ananya in improving family health outcomes in Bihar. Our approach to measuring program effects will be based on a comparison group design. While a comparison group design is generally not as strong as a well executed experimental design, given the implementation of Ananya other more rigorous methods are simply not feasible. The evaluation approach we propose later in this report allows us to factor in the dynamic health sector landscape of Bihar as we select our comparison group. In addition, our MLE approach acknowledges the important role of the local context and other nonhealth sector reforms (such as in education or transportation) that might influence family health outcomes. In addition, through the process evaluation

component of the MLE effort, we will carefully document and track the activities of other development partners in the health and related sectors to better assess the contribution of the Ananya program to overall health improvements in Bihar.

- ***The integrated nature of the family health approaches being implemented under the Ananya program will only allow us to measure the combined effects of the package of interventions and not of each component separately.*** The four grantees and their partners will be implementing a variety of coordinated and synergistic interventions simultaneously in the same set of districts in Bihar. As a result, our overall measurement approach will only be able to measure the combined effects of the package of interventions being implemented by the grantees under the program, and not separate effect of each component, with two exceptions. One exception is the WHP grant, which is using a randomized experiment to measure the effectiveness of its innovative private sector approach to improving infectious disease diagnosis and treatment. Another exception is the rigorous testing of two or three highly innovative family health solutions that will be implemented on top of the core package of interventions in select areas in the eight focus districts. These rigorous pilot tests will focus on innovations for which highly credible evidence is needed to inform scale-up decisions. To assess the effects of other interventions being implemented by grantees as part of the core package of interventions, we will examine changes in proximal outcomes closely related to specific activities or program components and rely largely on the process study component of the MLE effort. We will also exploit, to the extent possible, variations in program implementation to better understand the effects of specific activities.

- ***Broad scope of the Ananya program makes it challenging to obtain detailed information on all the areas of intervention as well as all target populations.*** As noted earlier, the Ananya program aims to affect processes and outcomes in a range of family health domains, including maternal and child health, reproductive health, nutrition, immunization, infectious diseases, and sanitation and hygiene, among others. Although it would be desirable to gather detailed information on the program's effects in each of these domains for all target populations and subpopulations, this will not be feasible due to time and cost constraints. Therefore, our approach was to work closely with the foundation and grantees to identify key focus areas of the program and specific grants and to prioritize the outcomes and target populations on which the MLE will focus, particularly for primary data collection. For instance, for our primary data collection, we will focus on women who have had a live birth in the past year, and for this sample, measure neonatal mortality and a variety of coverage indicators. For other target populations and measures, such as infant mortality rate, under 5 mortality rate, maternal mortality rate, and fertility rates among women in the reproductive age, we will rely on existing secondary data to the extent they are available. We will also examine data collected through grantees' project monitoring efforts, and will obtain qualitative data as part of the process evaluation to provide information on those areas that are not a major focus of the primary data collection.

- ***Ambitious outcome and impact goals and scale-up targets includes some risk that the grantees may not be able to attain the program targets as per the scheduled timeframe.*** The foundation has set very ambitious objectives and goals for the overall program and specific grants, which may or may not be achievable within the program's five-year life cycle. For instance, the grantees are expected to implement successful approaches at scale and achieve measurable results in the eight focus districts by the end of year 2 (that is, by the end of 2012). The project goals are that successful

approaches should get scaled up in the remaining 30 districts of Bihar by 2015. The MLE plan attempts to responds to the program parameters set by the foundation, as well as grantees' implementation plans, including when and how activities will be rolled out, and the time frame within which results are expected to be observable. For example, the proposed timing of the first follow-up data collection, which is scheduled for the end of 2013, allows additional time for scale-up (and changes in proximal indicators) to occur in the eight focus districts. Similarly, the impact analysis will be powered to detect ambitious but realistic changes in key outcomes during the five-year evaluation period.

- ***The MLE effort needs to be responsive to the diverse information needs of multiple stakeholders, and which may vary over time.*** A variety of stakeholders, including the foundation, grantees, the GoB, and various development partners are interested in the MLE for the Ananya program, and have specific needs. Through the various study components of the MLE plan, we will address the measurement and information needs of these stakeholders to the extent feasible. To ensure the greatest success of the MLE effort in meeting the stakeholders' needs, it will be critical to have strong communication and coordination between and among the MLE team and other internal and external stakeholders. To facilitate coordination, the foundation has created several mechanisms and forums for information exchanges among the various program partners in Bihar. These include quarterly partners' meetings, an MLE workgroup, a Project Coordination Committee (PCC), and a Project Action Committee (PAC). As we execute the MLE plan, we will be as responsive as possible given the basic design, and accommodate MLE needs of the program partners and external stakeholders as is feasible.

E. Road Map for the Rest of the Report

The remainder of the report is organized as follows. In Chapter II, we provide an overview of the MLE plan, including the conceptual framework underlying our overall approach, the key research questions to be addressed through the MLE effort, and an overview of the major evaluation components. Chapter III describes in more detail our proposed approach to conducting each of the major evaluation components. In Chapter IV, we summarize the data needs and sources for the impact evaluation component of the MLE plan, focusing on the primary data that we propose to collect. We conclude in Chapter V with an overview of the time line for major MLE activities, as well our reporting and dissemination plans.

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

II. CONCEPTUAL FRAMEWORK AND OVERVIEW OF THE MLE DESIGN

This chapter provides a brief overview of our proposed approach to MLE for the Ananya program. It begins with a discussion of the conceptual framework guiding our MLE approach, followed by a brief discussion of the research questions on which the MLE effort will focus and the broad study methods that we will use to address them. We provide a more detailed description of each of the evaluation components of the MLE plan in Chapter III.

A. Conceptual Framework

The conceptual framework underlying the MLE design (Figure II.1) illustrates the process by which the Ananya program expects to affect family health outcomes and ultimately achieve impacts. Through various innovative demand- and supply-side approaches implemented by the four grantees, the Ananya program aims to improve knowledge, attitudes, and social norms related to family health and to increase the availability, use, and quality of cost-effective family health interventions, leading to increased coverage and ultimately reduced mortality, fertility, and child malnutrition. In addition to the approaches implemented by Ananya grantees, the figure shows the various contextual, environmental, systems, and policy factors that will influence the implementation design and overall success of the program, as well as independently affect key family health outcomes, thereby highlighting the complexity and challenges of the MLE effort noted in the previous chapter. For example, changes in the coverage and quality of family health interventions are likely to be affected not only by the Ananya program, but by government and donor investments in the health sector through the NRHM and DFID-funded HSRP, as well as other health sector activities being conducted by donors and nongovernmental organizations (NGOs) throughout Bihar. In addition, interventions and improvements in other sectors, such as infrastructure, education, and water and sanitation, can affect maternal, neonatal, and child health outcomes. Finally, health outcomes and service utilization are affected by household-level factors including socioeconomic status, caste, and religion.

To understand the role of various contextual factors and external partners in observed changes in key outcomes and the ultimate impacts of the Bihar Initiative, the evaluation will pay considerable attention to the linkages between these factors and important proximal and intermediate outcomes over the program life cycle. In addition, as described below, the proposed MLE plan includes a rigorous and comprehensive process evaluation component, which will be used to document and assess changes in these factors over time, how they interacted with program implementation, and their affect on key outcomes.

B. Key Research Questions and Evaluation Components

As mentioned earlier, the MLE effort will focus on addressing three broad categories of research questions:

1. What family health approaches were implemented under Ananya? Did they achieve scale?
2. Did the implementation of these approaches at scale have impact? Were the highly innovative approaches tested by grantees effective?
3. What was the cost of implementing the program and were these costs effective?

Figure II.1. Conceptual Framework for MLE for the Bihar Initiative

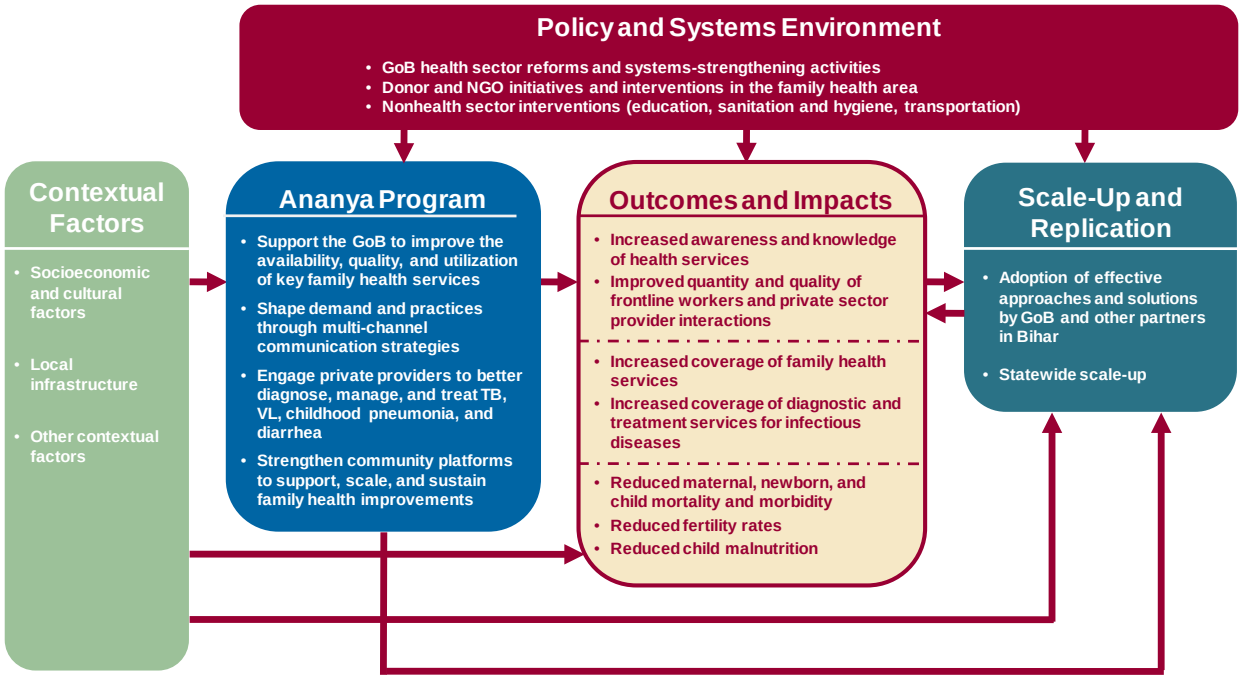


Table II.1 presents specific subquestions that the evaluation will examine in the course of addressing these broad sets of questions. The table also specifies the broad purposes of measurement and stakeholders for each question area and summarizes the approaches and data sources that we will use to answer the questions. As shown in the table, our investigation of these questions will feature a mix of qualitative and quantitative data sources and methods, which will improve the depth, rigor, and generalizability of our findings. For some questions, we will rely primarily on qualitative information and analysis, whereas for others the primary approach will incorporate quantitative data and methods. Most often, the two types of data and analyses will complement each other so that the MLE results will benefit from the specificity and rigor associated with quantitative methods and the explanatory richness and contextual value of the qualitative work. These questions will be addressed with a view to meeting the needs of a variety of stakeholders, including the GoB, foundation, grantees, and external stakeholders, as shown in the second column of Table II.1.

The MLE plan consists of three main evaluation components, each of which will inform (to varying degrees) our answers to the range of questions included in Table II.1: (1) process and scale-up analysis, (2) the impact analysis, and (3) cost and cost-effectiveness analysis. We briefly discuss each of these components next. A more detailed description of each component is provided in Chapter III.

1. Process and Scale-Up Analysis

The MLE effort will include a rigorous implementation or process evaluation, which will provide a comprehensive picture of what interventions are being implemented, how they are being implemented, and the factors affecting implementation. The purpose of this MLE component is to generate information to (1) understand implementation processes, successes, and failures that will inform program improvement; (2) inform our interpretation of impact analysis findings; and (3) guide replication of program innovations by GoB and other development partners.

Table II.1. Key Learning and Evaluation Questions

Evaluation Questions	Measurement Purpose	Approach	Primary Data Sources	Expected Outputs/ Deliverables
Documenting and Assessing Program Implementation and Scale-Up: The "How" and "Why" Questions				
What demand- and supply-side approaches to improving family health and infectious disease outcomes were implemented, how and where were they implemented, and at what cost? • How were the interventions rolled out? What geographic areas did they cover and how did program intensity vary across those areas? What were the systems and environmental contexts in these areas? • Were the interventions implemented as planned? What were the reasons for any deviations? • What were the roles of internal and external partners in implementing the interventions? What problems were encountered and how were they resolved? What challenges or barriers were common across all interventions and which were unique to specific interventions or approaches? What lessons were learned for future implementation? • To what extent did grantees integrate their approaches to improving family health outcomes? How did they collaborate, coordinate, and support one another during the course of the initiative? Were there key synergies between various approaches/interventions within and across grants and between grantees and external partners? • What are households' perceptions of the interventions? To what extent were they exposed to the interventions/aware of program activities? What were their perceptions of the services offered?	• Document implementation • Inform grantees' and foundation's decisions related to need for course corrections and program improvement – Provide guidelines and criteria for replication • Monitor and track progress – Monitor activities and outputs at grant level – Provide information for midcourse corrections • Inform GoB decision-making and the broader field – Understand process, successes, failures, and challenges of implementation and scale up • Understand interaction of grantees under a single initiative	• Develop process monitoring framework and indicators • Develop process monitoring MIS • Track achievement of and trends in outputs and key outcome indicators • Monitor policy-related developments and events • Conduct or draw from existing media content analysis • Rigorous process evaluations in years 2, 3, and 4, including field visits to select locations	• Program documents • Stakeholder interviews • Focus groups • Direct observation • Facility and frontline worker assessments • Project/grantee monitoring data • HMIS and ICDS HMIS data • Household survey data • Other administrative data and reports	• Feedback to the foundation, grantees, and GoB to support program monitoring and implementation efforts • Process/implementation reports in Q1 of 2013, 2014, and 2016
Did innovative family health solutions/approaches achieve scale in Bihar? To what extent and how did scale-up occur, and what factors enabled and inhibited scale-up? • To what extent did scale-up occur, both within the eight focus districts and across the state of Bihar? What were the planned mechanisms for scale-up for focus districts versus the state? What was the actual process through which scale-up occurred? What were reasons for deviations from plans, if any? • What are the key facilitators for scale-up of such integrated demand- and supply-side approaches to improving family health outcomes (knowledge, cost, collaboration, partnerships, political commitment, and so on)? What are key challenges and barriers? How did the Ananya program and particular grantees address various challenges to scale-up? • Did the GoB and private sector provide support (financial and nonfinancial) 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?	• Measure scale-up • Inform foundation and GoI decision making on replicating program approaches in other states • Inform GoB and the field – Understand successes and challenges of scale-up	• Process evaluation methods	• In-depth interviews with partners and external stakeholders • Surveys with key stakeholders • Administrative data	• Scale-up assessment reports in Q1 2014 and Q1 2016

Table II.1 *(continued)*

Evaluation Questions	Measurement Purpose	Approach	Primary Data Sources	Expected Outputs/ Deliverables
Measuring the Impacts of the Program and of Select, Highly Innovative Solutions				
Did the Ananya program contribute to improved family health and infectious disease outcomes in the eight focus districts and in the state of Bihar? • Did the program's integrated set of demand- and supply-side approaches improve key family health outcomes in the eight focus districts after two years? • Did the Ananya program succeed in improving knowledge, attitudes, and practices among consumers; the capabilities and performance of frontline workers and facility-based providers; and the availability and quality of essential family health and infectious disease services delivered in community and facility settings? • Did the combined efforts of the Bihar development partners to replicate and scale up successful family health interventions have an impact on key family health outcomes at the state level? How did the improvements in key outcomes across the state compare with improvements observed in the focus districts? • How did the effects of the program vary by key population subgroups?	• Measure overall contribution – Assess the effects of the Initiative's efforts in improving shorter term outcomes in the eight focus districts – Document changes in intermediate- and longer-term term outcomes • Inform foundation, grantee, and GoB decision making about scale up with Bihar • Inform the GoI and the broader field	Effects in eight focus districts after year 2: • Double-difference comparison group design • Estimate effects of package of demand- and supply-side interventions Statewide effects after year 5: • Pre-post design (no comparison group) • Assess adequacy of changes in key indicators • Dose-response and/or trend analyses, if possible	• Household, frontline worker, and facility-based surveys conducted at three time points: – Baseline: Fall 2011 – Midline: Fall 2013 – Endline: Fall 2015 • Available secondary data sources • Process evaluation data to interpret and explain impact evaluation results	• Baseline report in Q2 2012 • Midline report in Q2 2014 • Final report in Q2 2016
What highly-innovative value-added approaches to delivering family health services are effective? • Did the addition of innovative solutions such as ICT tools for frontline workers (such as mobile job aids and/or tracking systems) to the core integrated package of interventions implemented under Ananya lead to a greater improvement in key family health outcomes? • Did the addition of other innovative solutions (e.g., new or improved team-based incentive schemes) increase the effect of the initiative's core activities on key outcomes? • Were these interventions more effective for certain subgroups of the population than others?	• Demonstrate effectiveness of select, highly-innovative family health solution 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 – Disseminate results widely to encourage replication	Randomized controlled trial(s) • Identify select solution levers for rigorous pilot testing (TBD) • Determine appropriate unit of assignment (e.g., block, subcenter, or village) and sample size needs • Randomly assign units to treatment and control groups	• Baseline and follow-up surveys tailored to specific solution lever being tested • May involve household, frontline worker, and/or facility surveys • Timing of data collection TBD	• Report on pilot test results
Estimating Cost and Cost-Effectiveness				
• What was the overall cost of the Ananya program and the cost of broad program components? What was the unit cost of the program per beneficiary? What was the cost of the value-added or innovative solutions? • What were the proportional costs of major program components? What were the cost profiles for the various program components during the start-up, implementation maintenance, and scale-up phases of the project? • What was the cost-effectiveness of the program and specific components? What was the cost per beneficiary reached and per unit of desired outcome (for example, deaths or morbidity averted) for the program as a whole and for the value-added solutions? • What were the determinants of cost and cost-effectiveness? How did the unit costs and proportional costs of major components change with scaling-up? How did cost-effectiveness evolve over the course of the Ananya program, and to what extent was this driven by changes in unit costs versus changes in outcomes?	• Measure costs from foundation/program funder perspective • Assess effectiveness of program costs • Inform future replication of program approach or components by GoB, GoI, and other development partners and donors	• Develop tools to facilitate prospective collection of cost data from implementing grantees and partners, and to enable the allocation of costs to specific activities and/or program components • Collect financial/cost data from all internal and external partners, to extent possible • Leverage, as applicable, platform approaches being investigated as part of the Disease Control Priorities Project	• Work plans • Operations manuals • Reports • Budgets • Expenditure data • Training materials • Project/grantee monitoring data • Administrative data at the state, district, and block levels • Interviews with key staff from grantee organizations, the foundation, GoB, and other internal and external partners	• Costing study report in Q1 2013 and Q1 2016

ICDS = Integrated Child Development Services; ICT = information and communication technology; HMIS = health management information system; TBD = to be determined.

A key assumption of the Ananya program is that integrated solutions can be brought to scale and that delivery at scale of high-impact family health services and interventions will significantly reduce maternal, neonatal, and child mortality and morbidity. Given the importance to Ananya's success of achieving scale, a key component of the MLE effort will be to measure the extent to which scale-up occurred, understand and document the scale-up process, and identify the factors that facilitate and inhibit scale-up.

2. Measuring the Impacts of the Ananya Program and of Select Innovative Solutions

The ultimate goal of the Ananya program is to improve family health outcomes in Bihar. Therefore, measuring changes in key outcomes over the course of the program life cycle, and assessing the contribution of the program to these changes, is a primary component of our evaluation design. In particular, we will examine the effects of the overall package of interventions being implemented by the grantees in the eight focus districts and in the state overall. In addition to the set of demand- and supply-side interventions being implemented by grantees, some grantees will be piloting highly innovative solutions in select areas. As part of the evaluation, we will rigorously test the effectiveness of one to three of these select value-added interventions or approaches.

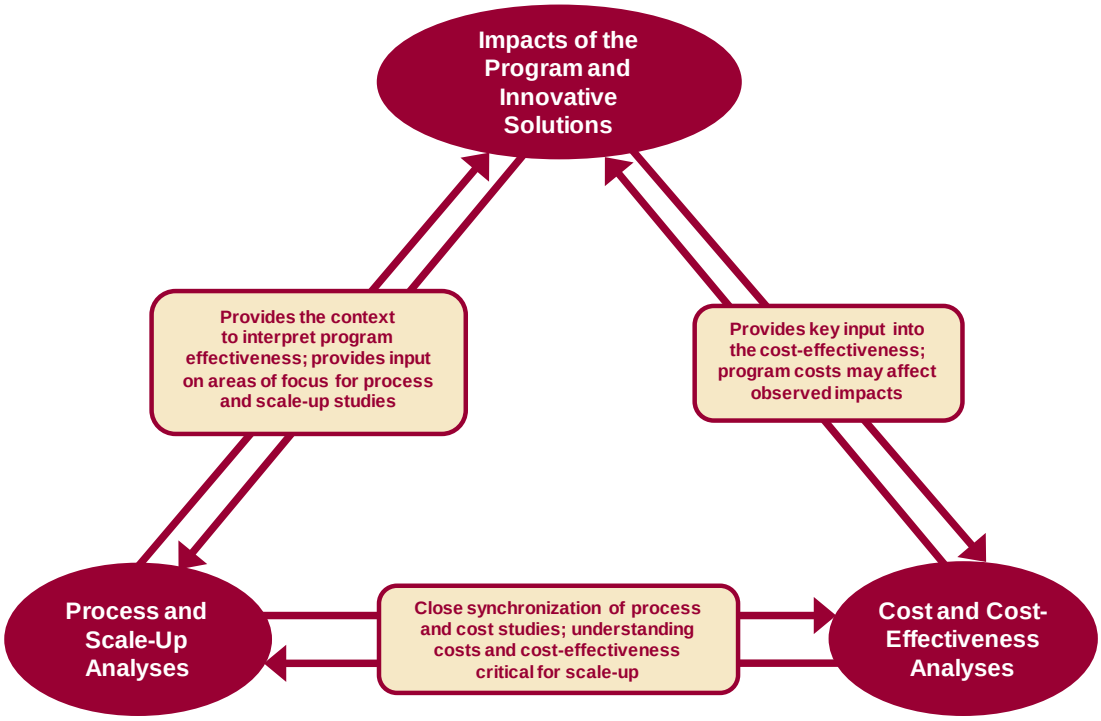
3. Cost and Cost-Effectiveness Analysis

Although evidence of effectiveness is critical for persuading development partners to adopt a new public health approach, the costs associated with replicating a program or approach and its impacts are also a key factor in replication and scale-up decisions. The costing component of our evaluation will aim to generate estimates of overall program costs, the costs of major program components, the key cost drivers, and replication costs. In addition, we will attempt to determine the cost-effectiveness of the program and the value-added solutions, and how cost-effectiveness may have evolved over the course of the Bihar Initiative.

We view the above three components of the MLE effort as very closely integrated and expect that each component will inform and provide input into the other two components (Figure II.2).

For example, the process and scale-up analysis will provide valuable information to better interpret the results of the impact analyses, by documenting and assessing what was implemented, the geographic coverage and intensity of various interventions, and the process through which families received various types of services and interventions. The process study will also provide critical data for the costing study by providing detailed information on the various program components and specific activities that were implemented, and the role of various actors in the program implementation process. Primary data collection for the impact analyses will provide important information on areas in which program implementation is working well or needs improvement, and thus will assist in identifying topics and focus areas for the process and scale-up study components. Finally, the impact analysis will provide critical information for the cost-effectiveness study.

Figure II.2. Interactions Among Key Study Components



III. ANALYTIC APPROACH TO THE MAIN EVALUATION COMPONENTS

In this chapter, we describe our approach to conducting each of the three evaluation components of the MLE plan. For each component, we start with a description of the overall objectives and questions that the study will address, followed by a description of the study design and data sources. We will further develop the approach for two of the study components—the process and scale-up analysis and the costing and cost-effectiveness analysis—after grantees implementation plans are refined and we have conducted an assessment of existing data sources. We expect to have memoranda outlining the final designs for each of these study components completed in spring 2012.

A. Process and Scale-Up Analysis

Overview and objectives. Knowing *what* interventions were implemented under the Ananya program, *how* they were implemented, and the *context* in which they are operating is essential for understanding how and why program impacts were or were not achieved and how successful approaches might be replicated and scaled up within Bihar and elsewhere. Documentation of the implementation process and associated challenges also provides information important for program improvement and accountability. The objectives of the scale-up analysis will be to understand the definition of and plans for scale-up, the extent to which it occurred, and the scale-up process. In addition, through the process and scale-up analyses, we will aim to identify the facilitating and inhibiting factors and document lessons learned for successful replication within Bihar and elsewhere.

Research questions. The process and scale-up analyses will focus on the following types of questions:

- ***What demand- and supply-side approaches to improving family health and infectious disease outcomes were implemented, and how and where were they implemented?*** How were grantees' interventions rolled out? What geographic areas did they cover and how did implementation intensity vary across those areas? In what environmental, systems, and community contexts were they implemented? Were the interventions implemented as planned and what were the reasons for any deviations?
- ***What were the roles of internal and external partners in the implementation process?*** What problems were encountered and how were they resolved? What challenges or barriers were common across all interventions and which were unique to specific interventions or approaches? What lessons were learned for future implementation?
- ***To what extent did grantees integrate their approaches to improving family health outcomes?*** What was the nature and extent of collaboration and coordination between and among grantees during the course of the program life cycle? Were there key synergies between various approaches/interventions within and across grants and between grantees and external partners?
- ***Did innovative family health solutions/approaches achieve scale in Bihar?*** How is scale defined in the Ananya context and what are the key indicators of scale-up to be measured at various time points? To what extent and how did scale-up occur? What were the planned mechanisms for scale-up within the eight focus districts and in

nonfocus districts across the rest of the state? What was the actual process through which scale-up occurred? What were reasons for deviations from plans, if any?

- ***What are the key facilitators for scale-up of integrated family health approaches?*** How do factors such as cost, collaboration, partnerships, political commitment and influence, and local context affect scale-up? What are key challenges and barriers? How did the Ananya program and particular grantees address various challenges to scale-up?
- ***What financial and nonfinancial support was provided by the GoB and other development partners for implementation and scale-up?*** How did the types and extent of support change over time? What role did such support play in achieving scale-up?

To obtain the depth of information necessary to answer these questions, we will focus on collecting data around several broad programmatic themes: (1) program/local context, (2) outreach and operations, (3) program staffing, (4) approach to service delivery, (5) collaboration and partnering among grantees, (6) partnership with the GoB, and (7) collaboration and partnering with other development partners and actors.

Data sources. The process and scale-up analyses will draw on primary and secondary data from a variety of sources:

- **Document reviews.** We will review documents produced by grantees during the planning and implementation stages of the project, including implementation progress reports, training manuals, and other such documents, as well as relevant documents produced or commissioned by the GoB and other development partners.
- **Stakeholder interviews.** A key feature of the Ananya program is the relatively large number of partners involved in and stakeholders affected by the program. In-depth or semistructured interviews with key stakeholders will provide insights into how the interventions are being implemented, the role of various actors in the implementation process, and what is working well and what is not. These interviews will also provide insight into the scale-up process. We plan to interview a broad range of partners and stakeholders, including foundation staff; implementing grantees; GoB staff at the state, district, and block levels; donors and NGOs investing in Bihar; community leaders; and private sector partners.[7]
- **Site/field visits.** In-depth field visits will provide the most detailed information on how program activities are being implemented and received on the ground. We will select eight districts—encompassing focus and nonfocus districts—in which to conduct site visits. In each district, we will visit two blocks, one subcenter per block, and two communities in each subcenter. During these visits, the field visit team will conduct interviews with district-level health officials, representatives of the grantee organizations, health facility staff, block-level administrative staff, frontline workers, and community

[7] We will also consider the approaches recommended by the foundation's diffusion and dissemination team on the recommended measurement methods. For example, we will assess whether social network analyses, which will include systematic collection of data from all key stakeholders, to quantitatively document the mechanisms of diffusion and spread is an approach that the foundation would want us to consider.

leaders. Another component of the field visits will be direct observation of facility operations, family–frontline worker interactions, and community dynamics. Observations of frontline workers' interactions can provide insight into the content and quality of those interactions, as well as frontline workers' knowledge and capabilities. The final component of the field visits will be to conduct focus groups with a range of community-level stakeholders, including community leaders, women of reproductive age, women with children aged 12 months or younger, husbands, and mothers-in-laws. Through these focus groups, we will obtain beneficiaries' perspectives on the rollout, quality, and benefit of various interventions and services. The advantages of focus groups are that they draw out group dynamics and provide a means of capturing the voices and personal experiences of beneficiaries that can provide rich insight into the perceptions of the interventions and its implementation that cannot be captured adequately by survey data.

- **Grantee implementation and tracking data.** Grantees' management information systems (MIS) and other reporting systems will provide information on the extent to which activities were implemented, how they were implemented, and the participation of relevant stakeholders in such activities.

- **Administrative data.** Administrative data can provide useful information on program implementation and scale-up. 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. Relevant administrative data, that are reasonable in quality, will be used as part of the process analysis.

- **Block-level tracking data.** As part of our process study, we will collect block-level data on a set of contextual variables and key health sector activities. This data will be collected in the roughly 340 blocks on which our primary data collection for the impact evaluation will focus; the data will provide input for both the implementation and impact analyses. Examples of contextual variables that will be tracked include block-level economic and sociodemographic characteristics; infrastructure; and the number, types, and locations of health facilities. Health sector activities include training and training content for frontline workers, skill labs to provide training to providers in facilities, and the distribution of supplies and materials to health workers. We plan to collect block-level tracking data systematically once each year, in order to provide sufficient information for the process analysis, as well as to be used for analytic modeling purposes.

Approach. We will use qualitative and quantitative data from the various sources mentioned above to document and assess the context in which program activities are being implemented, the implementation process, and the roles of various actors. More specifically, we will examine economic, demographic, cultural, and religious factors; the availability of structural and staff resources needed for program implementation and scale-up; and the financial and political support provided by GoB during different phases of the program. Our process analysis will also examine the fidelity of the interventions to the original plans and reasons for any deviations. For example, were the modifications or deviations a response to resource constraints and contextual factors, or did they evolve in the context of securing local ownership of the interventions? Are these deviations likely to affect the effectiveness of the program or specific components in improving family health outcomes? We will also attempt to understand the facilitators and inhibitors of successful implementation to draw lessons for future replication efforts.

An important component of the analysis will be to develop an operational definition of scale and identify key measures of scale up at various key program junctures in the program implementation process. We will work closely with the foundation's workgroup on diffusion and dissemination to adopt consistent set of definitions and measures. Using the recommended theoretical framework, we will examine the extent to which scale-up occurred and the mechanisms through which it occurred, as well as try to obtain a broad understanding of the fidelity of the interventions scaled across the state to what was implemented in the focus districts. For example, the analysis will examine whether explicit and intentional mechanisms were identified to facilitate scale-up, why those mechanisms were selected, the resource needs for those mechanisms, and whether they were successful. Deliberate mechanisms might include creating explicit linkages with other donors or GoB to integrate interventions, providing replication manuals, or placing program staff with other agency staff. Less-deliberate mechanisms might include publicizing the findings through various forums, such as meetings and workshops. The scale-up analysis will also investigate these same topics from the perspectives of the GoB and other development partners in Bihar, such as UNICEF and DFID, that might have been instrumental in the statewide scale up process (or that might have been aware of the interventions, but were not chosen for scale-up).

Our analysis of process and scale-up will draw on the conceptual framework underlying the Ananya program and logic models for each of the grants, as well as a theoretical framework that we will develop for understanding the dissemination process and mechanisms. Key elements of our analysis will include using more than one perspective and type of data to obtain information on key topics of interest and examining the consistency (or lack thereof) of information across various data sources. We will triangulate the qualitative data collected, assessing the extent to which multiple respondents and data sources provided similar inputs and insights and blending these with the quantitative data to provide a rich understanding of the different dimensions of program implementation and scale-up.

Timing of process and scale-up analyses. We will align the timing of the process evaluation with program implementation on the ground. We anticipate three rounds of process evaluation: fall 2012, fall 2013, and fall 2015. Each round will involve the collection and analysis of secondary data—such as grantee MIS or dashboard indicators compiled by the foundation, GoB's HMIS, and other administrative data—as well as field visits to select districts, blocks, and villages. The first round of process evaluation will occur roughly 12 to 15 months after implementation start-up for most grantees and will focus on characteristics of the interventions being implemented and how they are being rolled out; the context in which they are being rolled out; the acceptability, uptake, and perceived benefits of the interventions among beneficiaries and other stakeholders; and any initial bottlenecks or challenges. The subsequent rounds will additionally focus on implementation progress, including achievement of key outputs and outcomes and the extent to which scale-up is occurring. We will summarize the results of each round of the process and scale-up components in a report that aims to provide timely and actionable feedback to the grantees, the foundation, GoB, and other stakeholders. We will also use these findings as input into the reports on our impact analysis results.

B. Analysis of the Impacts of the Ananya Program and Its Innovative Elements

Overview and objectives. A key objective of the evaluation of the Ananya program is to assess whether the various approaches and solutions implemented under it were successful in improving family health outcomes. The GoB and foundation have set high targets for key coverage and impact indicators and it will be critical to measure the extent to which the program was able to reach these targets. It will also be important to assess whether the Ananya program was able to

improve equity in key health outcomes. The impact analysis will focus on measuring the effects of the integrated package of demand- and supply-side approaches being implemented by grantees and their partners, as well as the extent to which the program achieved its overall goals in terms of changes in key outcomes and impact indicators. In addition, we will attempt to rigorously test the effectiveness of select high-risk, high-reward solution levers that will be implemented on a pilot basis in select areas, on top of the core package of demand- and supply-side interventions.

Research questions. The impact evaluation component will seek answers to the following types of questions related to the overall impacts of the Ananya program and of select, highly innovative solutions:

- ***Did the program lead to improved family health outcomes in the eight focus districts and in the state of Bihar?*** Did the integrated set of demand- and supply-side approaches improve key family health outcomes in the eight focus districts? Did the combined efforts of the grantees and other partners to replicate and scale up successful family health interventions have an impact on key family health outcomes at the state level?
- ***Did the program succeed in increasing knowledge and changing attitudes and practices at the household level? Did it improve the capabilities and performance of frontline workers and facility-based providers?*** Did the program improve the availability and quality of essential family health and infectious disease services delivered in community and facility settings?
- ***How did the effects of the program vary by key population subgroups?*** Was the program able to reach more-marginalized populations, such as lower socioeconomic status groups and castes? What effects were observed for these subgroups relative to effects for women of higher castes or socioeconomic status?
- ***What innovative value-added approaches to delivering family health services are effective?*** For example, did the addition of innovative solutions (such as mobile job aids for frontline workers or mobile tracking systems) lead to enhanced effects of the program's core activities on key health outcomes? Did the addition of other innovative solutions such as new or improved incentive schemes lead to greater improvement in key family health outcomes?[8]

Next, we describe our planned approach to answering these questions. We first present our approach to assessing the effects of the overall Ananya program and then discuss our approach to measuring the effects of select, innovative solutions.

[8] Note that these examples are being used for illustrative purposes. We will work closely with the grantees and the foundation staff to identify which innovations should be rigorously tested.

1. Assessing the Overall Impacts of the Ananya Program

The implementation of the Ananya program will involve considerable collaboration, coordination, and synergies across grantees, particularly during the first two years of the program when implementation efforts will be concentrated in the eight focus districts. Because of the extensive overlap in target populations, geographic coverage, and key outcomes across the implementing grants, we will measure the combined effects of the package of interventions or activities being implemented by the grantees.

Our measurement of the overall effects of the Ananya program will be synchronized with the program's two planned stages of scale-up. During the first two years of the program (2011–2012), grantees will focus on implementing and scaling up their approaches and interventions in 8 focus districts in Bihar. Based on lessons learned from the implementation and scale-up process in the 8 districts, the grantees will support and facilitate the replication and scale-up of successful approaches in the remaining 30 of 38 districts of Bihar in years 3 through 5 (2013–2015). This two-stage scale-up approach suggests two key junctures in the five-year program cycle at which the success of the Ananya program should be assessed: (1) after scale-up has occurred in the 8 focus districts; and (2) at the end of the program life cycle, when statewide scale-up will have occurred and the overall effects of the program can be assessed.

Study design. Our approach to measuring the effects of the program at these two junctures attempts to maximize the potential rigor of the analysis while also taking into account the realities of grantee implementation plans and the contexts in which they are being implemented.

Assessing short-term impacts. A midline assessment of program effects will provide critical information on whether the integrated package of demand- and supply-side approaches being implemented in the eight focus districts are having an effect on key short-term outcomes. Evidence of success in the eight focus districts, combined with qualitative information on *how* interventions were implemented and practical lessons learned, will be critical to the program's efforts to promote, facilitate, and support the adoption of effective solutions into GoB policies and programs. In addition, evidence of effectiveness at this juncture will provide important information to grantees on whether and how to revise their implementation strategies to achieve program targets. We propose to conduct this assessment during year 3 of the program (in 2013), which will provide sufficient time after implementation in the eight focus districts to enable us to observe changes in proximal outcomes.

We will use a comparison group design to assess the short-term effects of the program in the eight focus districts. This approach improves upon a simple pre-post design by enabling us to distinguish program impacts from time trends or other concurrent factors that might influence outcomes. Having a comparison group is particularly important in the Bihar context, in which statewide development efforts by the GoB and other actors in the health sector could lead to improvements in health outcomes aside from the effects of the program. With a pre-post design, it would be difficult to ascertain whether (or to what extent) observed changes in outcomes in the eight focus districts are due to the Ananya program or to factors external to the program. A comparison group design would enable us to increase the *plausibility* of a causal effect of the initiative's efforts by offering a counterfactual representing what would have happened in the absence of the Ananya program. However, given the usual limitations of a comparison group designs, we cannot attribute observed effects to the Ananya program with certainty, since we cannot directly test the validity of the counterfactual.

For this approach to be credible, we must be able to identify comparison districts in Bihar that are similar to our focus districts in baseline outcomes and other key characteristics related to outcomes (such as demographic and socio-economic characteristics), and in which there are no intensive health sector activities by other donor partners. [9] We will use a mixed-methods approach to identify strong comparison districts. First, we will identify a set of potential comparison district matches using a statistical matching approach, drawing upon the most recent data available. (We hope to use district-level estimates published by the 2010 Annual Health Survey (AHS) as well as data from the 2010 census for the matching). We will supplement quantitative matching with qualitative information obtained from experts to assess the similarity of districts and to identify a final match for each focus district, taking into account the coverage and intensity of other development partners' activities in Bihar.

As noted earlier, one of the challenges we may face is in identifying comparison districts with relatively little to no health services intervention being implemented. In order to maximize our chances of getting a credible comparison group we will take advantage of the fact that we will be collecting data from all districts in Bihar in order to establish a statewide baseline. In particular, we propose to select comparison districts—based on baseline measures—just prior to the time of the midline assessment. Although we will still use baseline data to select matches, selecting the comparison districts in 2013 prior to when the mid-line survey is conducted, will allow us to take into account any changes in the scope and intensity of activities by other development partners in the various districts of Bihar. We will also consider alternative approaches to selecting the comparison districts. For example, one approach would be to use a weighted average of the full set of "appropriate" non-focus districts as comparisons.[10] More specifically, we could implement a kernel matching approach that computes weights based on the mean-squared difference between each focus and appropriate non-focus districts in baseline outcomes and characteristics related to outcomes. Alternatively, we could implement a synthetic control approach, which explicitly constructs a "synthetic" comparison district by choosing weights so that the synthetic district most closely resembles the focus districts at baseline (Abadie and Gardeazabal 2003, Abadie, Diamond and Hainmuller 2010).[11] These weighted approaches will be particularly useful if it proves difficult to find specific appropriate comparison districts that closely approximate the focus districts, and will also be useful as check of the sensitivity of the results.

---

[9] While a comparison group design is not as strong as a random assignment or regression discontinuity design, given the nature of the Ananya program implementation no other more rigorous design is feasible. Not attempting the comparison design feasible is also not an option, for otherwise we would not be in any position to make any statements about the plausible effects of the Ananya program, Our comparison design approach will carefully detail the exact methods used to conduct the matches, and we will conduct extensive sensitivity tests to assess the robustness of the results. If at the time of selection of the comparison districts it looks like this method simply will not work because there is absolutely no feasible comparison district(s), we will discuss the issue with the foundation team and provide a recommendation on the best course of action.

[10] We say "appropriate" nonfocus districts because if some *ex-ante* information is available suggesting that a non-focus district would be a particularly poor comparison (for example, due to intensive activity by other donors), then this district will likely be excluded from the "donor pool" before we start the matching process.

[11] In the synthetic control approach, the full set of appropriate non-focus districts is the "donor pool"; each district in the donor pool receives a weight that can be zero or positive, with the weights summing to one. Abadie, A. and J. Gardeazabal (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93 (1), 113-132. Abadie, A., A. Diamond, and J. Hainmueller (2010), "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105 (490), 493-505.

Assessing longer-term effects. Given plans to scale up successful approaches statewide, a comparison group design will not be feasible for the overall assessment of the program's success at the end of the current five-year program cycle. Because successful scale-up depends on the adoption and replication of effective approaches by GoB and other development partners, attribution of changes in outcome and impact indicators to the program alone are also not necessary. Instead, the purpose of this second-stage impact analysis is primarily to assess whether, at the end of the program cycle, the Ananya program met its objectives in terms of targeted changes in key indicators. Therefore, the focus will be on a pre-post design that measures changes in key outcomes among target populations between fall 2011 (the baseline) and fall 2015.

This basic pre-post design can be extended, depending on the availability of data, to provide suggestive evidence of the program's effectiveness. First, we can compare changes in outcomes in the 8 focus districts to changes in the 30 nonfocus districts or the state as a whole (a variant of a "dose-response" analysis). If we see greater pre-post changes in the 8 focus districts, where implementation of the Ananya program's approach will be more intense and of longer duration, this will suggest that the program did contribute to observed changes in outcomes at the state level. Second, we might be able to use existing secondary data, such as data from the District-Level Health Surveys (DLHSs) and AHS, to construct a (short) pre-period trend and, based on the assumption that the trend would have continued if the program had not been implemented, use this pre-trend as a counterfactual to assess program impacts.

Outcomes and data sources. The Ananya program aims to improve a broad range of demand- and supply-side outcomes across several family health domains. On the demand side, key outcomes include knowledge, attitudes, and practices related to maternal and newborn health, nutrition (with a focus on infant and young child feeding), and family planning (with a focus on post-partum use on intrauterine devices (IUDs) and tubal ligation). On the supply-side, outcomes include indicators related to the availability, quality, and content of family health services provided by frontline workers and at the facility level.

The ultimate goal of the program is to reduce maternal and childhood mortality; as a result, measurement of changes in mortality is a priority for the impact evaluation component of the MLE effort. We will rely on a combination of primary data sources and secondary data to address the overall needs of the evaluation and to assess changes in key impact indicators. Although it would be ideal to use existing secondary data sources to measure the impacts of the Ananya program, for several reasons, existing data will not meet the needs of the evaluation.[12] As a result, it will be critical to collect primary data from households, frontline workers, and providers at facilities. Our primary data collection will focus on the measurement of various coverage indicators and neonatal mortality impact indicator, while for other impact indicators such as infant mortality or under 5 child

[12] These reasons include (1) the fact that AHS household-level data (the most relevant source for household-level indicators) will not be released for research purposes, (2) the fact that the current AHS only covers the period through 2009 (and that the final round of AHS in which will be released in 2013 will provide data on household practices in 2011 which in reality will be a "baseline" for the Ananya grantees), and (3) the dearth of information available through secondary data sources on several key outcome domains, including family–frontline worker interactions and frontline worker capabilities, motivation, and performance.

mortality, maternal mortality, contraceptive prevalence rate and total fertility rates, we will rely on secondary data sources to the extent they are available.[13]

The reasons for focusing on neonatal mortality in our primary data collection are two-fold. First, neonatal deaths account for the majority of all infant deaths and deaths to children under five years of age in India. Second, the focus of BBC and CARE interventions is largely on the period between the last trimester of pregnancy and the first month of life. These grantees are focusing on factors that are more likely to affect outcomes and survival during the neonatal period, including birth preparedness, the quality of delivery and immediate postpartum care, infection control, prevention and management of hypothermia, and immediate and exclusive breastfeeding.

While we considered measuring infant mortality as part of our primary data collection, our intended target population of women who gave birth in the previous 12 months would not allow us to accurately capture infant mortality, since the children of these mothers would not have been exposed to a full year of potential life. In order to measure infant mortality accurately, we would have to extend our reference period by another 12 months, which would likely lead to substantial recall error for other key outcomes of interest to the program. As a result, we will rely on secondary data sources for measuring IMR.

Obtaining primary data on other indicators, such as maternal mortality, poses significant challenges due to the low frequency with which maternal deaths occur—making it a rare event requiring very large sample sizes to measure—and due to high levels of misclassification and underreporting. Given the difficulty and cost of measuring maternal mortality[14] we propose to rely on secondary sources such as the AHS or SRS to track changes in this indicator. Similarly for other impact indicators such as under 5 mortality, total fertility rate and contraceptive prevalence rates, we will rely on secondary data that may be collected in the future.[15]

In Chapter IV, we provide a more detailed description of potential existing data sources and information that will be available for our analysis, and the types of primary data we propose to collect for the impact analyses.

Sample design. To answer the key study questions about the short- and longer-term impacts of the Ananya program, we will conduct baseline (2011), midline (2013), and endline (2015) surveys with households, frontline workers, and facilities (primary health centers [PHCs], including providers). As we consider the appropriate sampling approach and sample sizes for the baseline data collection, we have the following three goals in mind:

---

[13] We also intend to measure the effect of the Ananya program on stillbirths, primarily through the household listing which we will conduct as part of our household survey effort (as discussed in Section IV).

[14] We explored the possibility of examining maternal "near miss" (or, severe acute maternal morbidity (SAMM)) events to measure and assess changes in the quality of facility-based obstetric care. However, this measurement approach is more valuable in contexts where high-quality service use data is collected systematically at the facility level and in facilities that regularly provide emergency obstetric care. Since grantee's activities at the facility level will focus largely on PHCs, which typically refer patients with severe obstetric complications to community health centers (CHC) or district hospitals, this measurement approach will likely not be effective at measuring the effects of grantees' interventions.

[15] We will also draw on COHESIVE's impact evaluation of the World Health Partners grant to describe the impact of the program on key infectious disease outcomes not measured through the impact analysis conducted by Mathematica and PHFI.

1. To measure the success of the Ananya program by comparing key health outcomes at baseline and endline across the state
2. To measure the success of the program in the eight focus districts by examining changes in key outcomes between baseline and midline relative to a set of comparison districts
3. To obtain district-level estimates for some key outcomes, particularly in the eight focus districts, to inform the grantees' decision making and help them reestablish benchmarks and targets as necessary

Having carefully assessed several potential sampling approaches and considering the advantages and disadvantages of each, we propose implementing surveys in all 38 districts in Bihar.[16] Within each district, we propose a three-stage sampling strategy. In the first stage, we will select a representative sample of blocks (the primary sampling unit, or PSU) in the district; in the second stage we will draw a representative set of secondary sampling units (SSUs) in the sampled PSUs.[17] In the final stage, we will conduct a household listing in each sampled SSU to identify households containing the main population of interest: women who gave birth in the past 12 months. Among eligible women, we would use systematic sampling to obtain the desired number of women in each SSU. This sampling approach would enable us to meet the goals described earlier and it has the following features:

- It will enable us to obtain district-level estimates for some key outcomes in *all* districts. These estimates will be of interest to grantees, would help to increase GoB ownership of the data, and will more generally allow for broader uses of the data beyond our evaluation purposes.
- It will provide us some flexibility in selecting districts for our comparison group design based on current information on health service activities in other districts in Bihar. Selecting comparison districts now carries some risk that they might not end up being appropriate comparisons if some development partners ends up providing similar services as the Ananya program. Because our sampling approach yields a representative sample in each district, the full set of districts in Bihar will be available to us as comparison district options, and in addition, we can use the baseline data collected to help with the district matches.
- Selecting blocks in the first stage will make it easier to obtain contextual and programmatic data at the block level for a well-defined number of blocks, which then can be used for analytic and modeling purposes for our impact analysis.

For the frontline worker surveys, we propose sampling all the anganwadi workers (AWWs), accredited social health activists (ASHAs), and auxiliary nurse midwives (ANMs) attached to the

[16] We also considered a sampling approach which involved creating and sampling from three strata: (1) the 8 focus districts, (2) a set of comparison districts and (3) the remaining 24 or so districts; the samples from these strata could be appropriately weighted and combined to provide state level estimates. However, our calculations suggested that this sampling approach leads to very large design effects and is therefore quite inefficient relative to sampling in all districts. Sampling in all districts also has additional advantages, as noted in this section.

[17] Following previous surveys in India, we will define SSUs as villages in rural areas and as Census Enumeration Blocks (CEBs) in urban areas.

sampled SSUs. Surveying frontline workers from the same communities has the advantage that the additional costs of the surveys are likely to be low, because the data collection team will be in the SSUs for the household survey. In addition, this sampling approach could enable us to conduct some correlational analyses that relate village-level household health outcomes to the knowledge, attitudes, and practices of frontline workers in these sampled communities. For the facility/provider surveys, we propose surveying the block PHC administrators and the providers in the PHC's sampled blocks. This approach will provide a representative sample of block PHCs and PHC providers at the state level.

Sample sizes. It is important to ensure that the sample for the household survey is sufficiently large to statistically detect impacts of the desired size across the various research designs. We computed the minimum detectable impact (MDI)—the smallest impact that can be statistically distinguished from zero with high probability—for a range of possible sample sizes. We did this by systematically varying the components of the sample size—the number of blocks, SSUs, and respondents—in order to determine which combination yielded acceptable MDIs for the lowest total sample size. As we calculated the MDIs, we were aware of the importance of estimating impacts for a wide variety of outcomes with different baseline prevalence rates. This includes the neonatal mortality rate (NMR) which, as noted above, is a crucial outcome that the Ananya program hopes to affect but one that has a low prevalence rate and, therefore, requires a larger sample size to detect expected impacts compared to other outcomes that we will measure.[18] We were also cognizant of the need to have sufficient power to be able to perform subgroup analyses on some key coverage indicators, particularly for some key subgroups such as caste and socioeconomic status.

Based on our analysis, we propose to select nine blocks in each district, three SSUs per block, and an average of 15 households per SSU.[19] A few districts have fewer than nine blocks; here we will select all blocks and adjust upwards the number of SSUs. This yields a sample size of 15,390 households. As seen in Table III.1, for the statewide pre-post comparison, this will yield an MDI of about 6 per 1,000 for the NMR (from a projected baseline of 32 per 1,000).[20] As the table shows, the sample sizes we propose will enable us to detect improvements in most other key coverage indicators.

This proposed sample size will also give us sufficient power to detect impacts for several subgroups of interest to the study. We will identify and estimate subgroup impacts for groups that are either being specifically targeted by the grants (such as members of marginalized populations), or

[18] The Institute for Health Metrics and Evaluation team is working on producing district-level estimates of child mortality. We will discuss its approach and methodology and whether it would be feasible to generate such measures for Bihar.

[19] The crude birth rate in Bihar was about 28 per 1,000 in 2009 according to SRS data. Since the average household size is about 5.5, we will have about 28 births in the past year in 180 households. This implies that roughly 16 percent of households would have had a birth in the past year, and suggests that we will require at least 95 households in an SSU to obtain our proposed sample sizes. (SSUs in urban areas will have to be larger as these tend to have lower birth rates than in rural areas). We will also combine small SSUs as necessary.

[20] The NMR from the 2005–2006 NFHS is 40 per 1,000 live births. These numbers were 35 per 1,000 according to the first round of AHS, which gathered information in 2007-2009. Assuming these rates continue to experience a similar trend we project the current NMR to be about 32 per 1,000.

Table III.1. Minimum Detectable Impacts (MDIs) for Assessing Impacts

Design and Sample	Total Districts	Total Respondents per Round[a]	Baseline Prevalence: NMR (Per 1,000 Live Births) 32	Baseline Prevalence: Various Coverage Outcomes (Percentage Points) 10[b]	20[c]	30[d]	40[e]
				MDI (Percentage Points)			
Pre-Post Design							
Women Who Gave Birth in Past Year							
All	38	15,390	6	1.3	1.7	1.9	2.0
50 percent subgroup	38	7,695	8	1.5	2.0	2.3	2.4
33 percent subgroup	38	5,130	9	1.7	2.2	2.5	2.7
20 percent subgroup	38	3,078	12	2.0	2.7	3.1	3.3
Frontline Workers	38	2,052	NA	2.4	3.1	3.6	3.8
Comparison Group Design[f]							
Women Who Gave Birth in Past Year							
All	14	5,670	--[g]	4.2	5.5	6.3	6.8
50 percent subgroup	14	2,835	--[g]	4.9	6.5	7.5	8.0
33 percent subgroup	14	1,890	--[g]	5.6	7.4	8.5	9.1
20 percent subgroup	14	1,134	--[g]	6.7	8.9	10.2	10.9
Frontline Workers	14	756	NA	7.8	10.5	12.0	12.8
				Sampling Error (Percentage Points)			
District-Level Estimates							
Individuals[a]	1	405	--[g]	2.6	3.5	4.0	4.3

Notes: Calculations are for binary outcomes and assume the following: intraclass correlations (ICCs) of 0.08 at the SSU level, 0.02 at the block level, and 0.04 at the district level for coverage outcomes (the mean ICCs for a variety of coverage indicators from the DLHS-3); ICCs of close to zero at the district and block levels and 0.04 at the SSU level for NMR (from the DLHS-3) ;an R^2 of 0.3 at the individual, SSU, block, and district levels; a correlation of 0.3 between outcomes at baseline and follow-up in the sampled districts, blocks, and SSUs; and a response rate of 90 percent. The MDIs for the NMR are rounded up to the nearest whole number. All of these calculations assume a two-tailed test with 80 percent power and 5 percent significance.

[a]Assumes 9 blocks per district; 3 SSUs per block; and 15 respondents and 2 frontline workers per SSU.

[b]For example, percentage having a post-natal checkup at home within 48 hours (7.9 percent).

[c]For example, percentage initiating breastfeeding within one hour when a skilled attendant is present (19.2 percent).

[d]For example, percentage having at least 3 ante-natal checkups (26.3 percent); percentage of safe deliveries (31.9 percent); percentage using contraceptives (33.3 percent).

[e]For example, percentage with knowledge of 2 or more newborn danger signs (61.6 percent); fully immunized children (41.4 percent).

[f]Assumes 8 focus districts and 6 matched comparison districts.

[g]We do not plan to assess the impacts on NMR in the comparison group design as we would not expect to see large differences in NMR within the two-year timeframe of this analysis.

NA = not applicable.

where we expect impacts might differ by the characteristic (such as parity).[21] Examples of subgroups of interest include members of scheduled castes/scheduled tribes (23 percent of the population), women in the poorest socio-economic quintile in the state (20 percent of the population), and

[21] As part of our analysis, we will estimate impacts separately for each relevant subgroup where relevant and feasible, and will also test whether the estimated impacts differ significantly across subgroups defined by a given characteristic.

women who have recently given birth to their first child (24 percent of women). Table III.1 suggests that we will have sufficient power to detect subgroup impacts of under 4 percentage points for key coverage outcomes, even for small subgroups comprising only 20 percent of the population (such as members of SCs/STs).

Table III.1 also shows the MDIs for the comparison group design assuming that we are able to select six comparison districts.[22] These MDIs are higher but are still well within the goals of the program for some key coverage indicators. The table also shows the MDIs for frontline worker surveys (assuming a single AWW and ASHA per sampled community). Although the MDIs for frontline workers are considerably higher than for the household sample, we expect they are more than sufficient given the large changes in frontline worker outcomes that we expect. Finally, the table shows the sampling error rates that would be generated by these sample sizes for providing district-level estimates.

2. Assessing the Effects of Selected Highly Innovative Family Health Solutions

As described earlier, our measurement of the effects of the Ananya program will include the rigorous testing of a few highly innovative solutions being implemented by grantees. Innovative solutions that have been recommended for rigorous testing include the use of mobile technology to improve the provision of care by frontline workers, using team-based incentives at the sub-center level to improve the quality of services provided by frontline workers, and the introduction and promotion of micronutrient power for children and postnatal mothers.[23] These innovative approaches are likely to be cost-intensive and have no precedent in Bihar, which may suggest that rigorous evidence would be needed to determine the cost-effectiveness of these approaches before promotion of widespread scale-up.

We expect the MLE effort for the Ananya program to include rigorous testing of one to three highly-innovative solutions to be implemented under the IFHI, Shaping Demand and Practices, and/or community mobilization grants. At the time that this report was written, each grantee (with the exception of World Health Partners) was in the process of developing a concept note describing the innovative solutions that it plans to implement, highlighting the ones that should be considered for rigorous testing. Based on these concept notes and other input from the grantees, the foundation will select one to three innovations to be rigorously tested by Mathematica and its partners. The kinds of criteria being considered for selection of specific innovations for rigorous testing include:

[22] Although we will match each focus district to a single comparison district, some focus districts could be matched to the same comparison district. Therefore the number of comparison districts will likely be fewer than eight; we assume that there will be six of these.

[23] As mentioned earlier, COHESIVE-India is conducting a rigorous evaluation of the WHP grant, which is testing a social franchising model that includes the establishment of new telemedicine centers on infectious disease outcomes. Although WHP will also be implementing initially in the eight focus districts along with the other three grantees, its implementation and scale-up plans are relatively independent of those of the other grantees and allow for experimental testing of WHP's approach.

- Highly credible evidence is required to motivate scale-up by GoB and external partners and stakeholders.
- The specific innovative solution is expected to have a very large impact, and has a realistic chance of proving sufficiently successful and cost-effective in a short period, to prompt scale-up by external stakeholders.
- Rigorous evaluation (for example, a randomized design) is feasible and potential contamination of the control group is minimal.
- The available sample size is sufficient to ensure adequate statistical power to detect effects, which might be a particular concern when randomization can take place only at a relatively high level (such as the block level), or when the intervention targets small subgroups of the population.

Study design. Our approach to measuring the effectiveness of the innovative solutions will be to use a rigorous random assignment design. For several of the interventions being proposed for rigorous testing, it seems feasible to have subcenters be the primary unit at which the intervention takes place, and so to randomize subcenters into treatment and control groups. From a purely statistical perspective it is generally desirable to randomize at the lowest level feasible (say a village as opposed to a block), because this results in greater statistical power. However, programmatic considerations are also important, in particular the level at which a particular solution will likely be delivered in practice. Potential contamination of treatment and control groups is a further concern. For example, if an intervention is randomized at the village level but delivered at the subcenter level, treatment and control frontline workers who are linked to the same subcenter could talk about or share their tools or practices, which would make it difficult to evaluate impacts. Based on recent discussions with the grantees, we anticipate it should be feasible to randomize at the subcenter levels.

Data source. In order to assess the impacts of the innovative solutions, we will have to collect data from a random sample of households and/or frontline workers in the treatment and control areas. If the ICT or team based incentive innovations are selected by the foundation for rigorous testing, it is likely that we will be able to leverage the baseline data collected as part of the overall impact evaluation. For example 2 to 3 blocks may be selected for each intervention to be tested, and subcenters randomly assigned to treatment or control status. In the case of each of these interventions, leveraging the baseline data would be feasible as the proximal outcomes collected as part of the baseline are largely relevant for these interventions. Hence, baseline data could be collected on additional sample members in subcenters in the selected blocks to form the baseline for the experimental study (and if needed, we can tweak to the current baseline instruments to capture any additional variables of interest relevant to the experiment).[24] We will collect follow-up data targeted on key outcomes and target populations of interest about a year to 18 months after the implementation of the intervention begins.[25]

[24] The MMP innovation will need a very different type of sample and baseline variables, and we do not anticipate that it can be relatively easily folded into our current baseline measurement, unlike the other two proposed innovations.

[25] Other crucial aspects of the data collection are documenting implementation of the solution lever and collecting data on costs. Detailed information on implementation is essential to interpret the quantitative impact estimates from any experiment. In addition, detailed data on costs are essential to conduct a cost-benefit analysis and to evaluate whether the solution lever is likely to be cost-effective at scale.

Sample sizes. Table III.2 presents sample sizes and MDIs for a range of baseline prevalence rates of indicators. The table shows that, for the proposed sample size, we would be able to detect a roughly 4 to 8 percentage-point change in key outcomes as a result of the innovation. (The actual indicators to be measured will depend on the innovative solution being tested). Given the short time frame of this component of the evaluation, it is unlikely that we will consider low-prevalence mortality outcomes for these assessments.

In order to determine the appropriate sample size, we systematically investigated the response of the MDIs to the number of subcenters and the number of respondents per subcenter. We recommend that approximately 60 subcenters be randomized into equal treatment and control groups, and that one to two villages in the catchment area of each subcenter be randomly selected in the case of subcenter-level randomization, and that 25 respondents in each subcenter be surveyed. This will enable us to detect an impact of between 6.3 and 10.5 percentage points for the main sample and impacts of between 7.3 and 12.2 for a subgroup of half of the population. We will be able to detect an impact of between 8.0 and 13.3 percentage points for the frontline worker sample. For fewer subcenters, the MDI's are slightly larger.

Table III.2. Minimum Detectable Impacts for Experimental Evaluations of Innovative Solutions

Unit	Sample	Number of Units (Treatment)	Number of Units (Control)	Number of Units (Total)	Number of Respondents per Round (Total)[a]	Baseline Prevalence (Percentage Points)				
						10	20	30	40	50
						Minimum Detectable Impact (MDI)				
Subcenter[b]	Women in the Target Population									
	All	30	30	60	1,500	6.3	8.4	9.6	10.3	10.5
	50 percent subgroup	30	30	60	750	7.3	9.7	11.1	11.9	12.2
	33 percent subgroup	30	30	60	500	8.2	10.9	12.5	13.3	13.6
	Frontline Workers	30	30	60	540	8.0	10.6	12.2	13.0	13.3
Subcenter[b]	Women in the Target Population									
	All	25	25	50	1,250	6.9	9.2	10.6	11.3	11.5
	50 percent subgroup	25	25	50	625	8.0	10.7	12.2	13.1	13.3
	33 percent subgroup	25	25	50	417	8.9	11.9	13.7	14.6	14.9
	Frontline Workers	25	25	50	450	8.7	11.7	13.4	14.3	14.6
Subcenter[b]	Women in the Target Population									
	All	20	20	40	1,000	7.7	10.3	11.8	12.6	12.9
	50 percent subgroup	20	20	40	500	8.9	11.9	13.7	14.6	14.9
	33 percent subgroup	20	20	40	333	10.0	13.3	15.3	16.3	16.7
	Frontline Workers	20	20	40	360	9.8	13.0	14.9	16.0	16.3

Notes: MDIs assume an R^2 of 0.3 at the individual, village, and block levels; an intraclass correlation (ICC) of 0.08 at the village level and 0.02 at the block level; and a survey response rate of 90 percent. The calculations assume a two-tailed test with 80 percent power and 5 percent significance.

[a]Assumes 25 respondents sampled per subcenter and 9 frontline workers sampled per subcenter.

[b]Assumes one or two villages sampled per subcenter.

C. Cost and Cost-Effectiveness Analysis

Overview and objectives. Although evidence of effectiveness is critical in persuading development partners to adopt a new public health approach, the cost associated with replicating a program or approach is also a key factor in replication and scale-up decisions. The costing component of the MLE design will provide the foundation, GoB, and other development partners

with information on the overall cost and cost-effectiveness of the Ananya program. More specifically, it will aim to generate estimates of overall program costs and the costs of broad program components and phases (start-up, implementation, maintenance, and scale-up. It will also attempt to determine the cost-effectiveness of the program and the innovations selected for rigorous testing. The costing study will be conducted from the perspective of the foundation (or program funder), and will not attempt to capture all economic costs, such as household or societal costs associated with program implementation and uptake of services delivered or enhanced under the program.

Research questions. The cost and cost-effectiveness analysis will address the following types of questions:

- ***What was the total cost of the Ananya program and of the broad program components?*** What was the unit cost of the program per beneficiary? What was the unit cost for community- and facility-level and cross-cutting interventions, and for public and private sector interventions? What was the cost of the value-added or innovative solutions?
- ***Within each intervention category, what were the proportional costs of major components?*** How were costs allocated across personnel, supplies, equipment, or other cost categories? What were the cost profiles for the various program components during the start-up, implementation, maintenance, and scale-up phases of the project?
- ***What was the cost-effectiveness of interventions and the overall program?*** What was the cost per beneficiary reached and per unit of desired outcome (for example, deaths averted, morbidity averted, or disability-adjusted life years saved) for the program as a whole and for the value-added solutions?
- ***What were the determinants of cost and cost-effectiveness?*** How did the unit and proportional costs of major components change with scale-up? How did geographic and contextual variables influence cost-effectiveness over the course of the program?

Approach. To answer these research questions, we will adopt the widely used activity-based costing (ABC) methodology and use what is referred to as an "ingredients" approach to cost estimation (Edejer et al. 2003; Chee and Makinen 2003; Fiedler 2003). The ABC methodology involves the assignment of costs (direct and indirect) to specific activities or categories of activities. This assignment process uses a combination of (1) direct tracing of costs to activities and (2) apportioning of costs to activities (in cases in which costs have to be allocated to multiple activities). This method tends to provide actionable cost information from a policy and program perspective because it provides cost estimates for specific program activities (or interventions) and allows for the inclusion of both direct and indirect costs.

The starting point for the ABC method is the development of a detailed description of the program and identification of broad activities that will be conducted as part of each program component (referred to as cost centers). For the ABC method to be successful, the list of activity areas must be exhaustive for a given time period to avoid double-counting of any resources used in implementing the program. Specification of the time period is a key element of the ABC approach and facilitates calculation of activity costs during different program phases (such as start-up, maintenance, or scale-up). Examples of activities include formative research, the development of media messages and radio broadcasts, frontline worker trainings and supervisory visits, and planning and coordination, among others. We will give consideration to the categorization of program

components and activities; different categorization schemes might be needed to answer different questions, thereby requiring a flexible approach to the collection of cost data.

The next step in the proposed costing approach involves identification of the inputs used for each activity specified in the first step. Ideally, we would collect information on both the quantity and price of inputs used for each activity, both for transparency purposes and to maximize the utility of the study results to stakeholders in Bihar and elsewhere. However, this may not be possible due to data constraints that may require broad categorizations of inputs. Inputs will include both capital and recurrent costs, such as personnel time, building space and equipment, materials, and overhead costs. Some (direct) costs will be directly attributable to a specific activity—such as transportation, meals, and per diem for a particular training. Other costs, such as staff salaries, might be shared across many activities within a given program. These costs will be allocated to specific program activities using various methods. For example, allocation of personnel time can be based on an analysis of timesheets, interviews, and/or administrative data. Other indirect costs will be allocated based on overall recurrent cost percentages. We will use the sum of direct and indirect costs for all program activities to estimate the program's total costs for a given period.

The cost-effectiveness analysis will extend beyond the costing study and compare program costs with program outcomes. This additional step requires allocating the costs of each activity toward specified outcomes, either based on quantitative methods or qualitative discussions with staff. This final costing exercise will provide data for the numerator of the cost-effectiveness ratio: the ratio of the cost of the program to the impacts of the program. The outcome measure in the denominator of these ratios (the impacts) will be based on the baseline, midline, and endline data collected as part of the impact evaluation component of the MLE effort. The two broad types of cost-effectiveness ratios that we will estimate are (1) cost per unit of beneficiary reached (for example, cost per newborn or infant reached) and (2) cost per unit of impact (for example, cost per new acceptor of postpartum family planning or cost per maternal/infant death averted).

To the extent possible, we will attempt to conduct a full costing of the program, including both financial and economic costs, from the perspective of the foundation. Financial costs are incurred costs, or actual expenditures. Economic costs include the value of all items used to implement the program, including unpaid items such as donated goods and volunteer time. We will attempt to capture all costs of implementing program interventions, including costs incurred by the foundation, implementing grantees, GoB, other key development partners, and households.

Data sources. We will use both retrospective and prospective data from a variety of sources for the costing study. Key data sources will include the following:

- Financial records used for routine financial and management reporting by the foundation, implementing grantees, and, to the extent feasible, other partners
- Program MIS data generated as part of grantees' project monitoring efforts
- Key informant interviews with foundation, grantee, GoB, and other partner organizations' staff
- Household, frontline worker, and facility surveys collected as part of the impact evaluation
- Other administrative data, including staff records, HMIS, Integrated Child Development Services' (ICDS) MIS, and various programming data

Some aspects of data collection for the costing study will be embedded in the impact and process evaluation data collection. For example, if feasible, we will attempt to collect information on costs to households of accessing various interventions (such as out-of-pocket expenditures or opportunity costs) through the households surveys conducted as part of the impact evaluation. We will collect information on time allocation of facility-based staff and frontline workers through interviews conducted as part of the process evaluation study. In addition, the MLE team plans to develop tools to facilitate prospective collection of cost data from implementing grantees and partners, and to enable the allocation of costs to specific activities and/or program components. We will develop these tools as the grantees refine their program components and specific activities. To the extent possible, MLE will also leverage the platform approaches being investigated as part of the Disease Control Priorities Project to capture data on external partners' costs and outputs.

Challenges. Several challenges to estimating the costs and cost-effectiveness of the Ananya program might require us to modify the costing approach proposed here:

- **Compiling complete, accurate, and appropriately disaggregated internal partner costs.** The Ananya program will be implemented by four grantee consortiums comprised of multiple NGOs, each of which will have its own accounting system and cost structure. As a result, it might be difficult to collect comprehensive and comparable data across all implementing organizations. In addition, GoB is a key partner of the Ananya program, and the relevant government departments might not have a cost accounting system capable of quantifying all of the types and quantities of resources used to support the program's activities.
- **Collecting cost data from external partners.** Successful scale-up of the Ananya program will rely on the collaborative efforts and investments of multiple external (public and private sector) partners, such as UNICEF, DFID, and mass media vendors. Accessing comprehensive or disaggregated cost or budgetary data from these organizations might not be possible, making accurate cost estimation difficult.
- **The scale and integrated nature of the Ananya program.** The Ananya program is expected to cover a large geographic area and encompasses many interrelated components aiming to affect similar outcomes, adding complexity to the costing study and making it challenging to allocate costs to specific outputs and outcomes for the cost-effectiveness analysis.
- **Determining which costs and benefits are part of the program.** As mentioned previously, many development agencies operate in the health and other sectors in Bihar. The activities conducted and services provided by these agencies could have direct or indirect effects on the implementation and impacts of the Ananya program, making allocation of costs and benefits of the program challenging.

IV. SURVEY DATA FOR THE IMPACT ANALYSIS OF THE ANANYA PROGRAM

Chapter III described our approach to assessing the overall impacts of the Ananya program. This chapter provides an overview of the data needs and sources for the impact analysis. We start with a brief description of relevant secondary data, followed by a description of the primary data that we propose to collect for this study.

A. Existing Household Survey Data

In the initial planning stages for the MLE component, the foundation anticipated that the impact evaluation of the Ananya program would be based largely on data from existing large-scale survey efforts in India, namely the NFHS, DLHS, and AHS. However, due to the timing of these existing surveys, as well as uncertainty about future data collection efforts, it was determined that additional primary data collection would be required to ensure a credible and comprehensive impact assessment of the program. In addition, although these surveys contain a range of relevant coverage measures, they include very limited measures of key proximal outcomes—such as knowledge, attitudes, and self-efficacy—and almost no information on family–frontline worker interactions, a key outcome domain for the Ananya program.

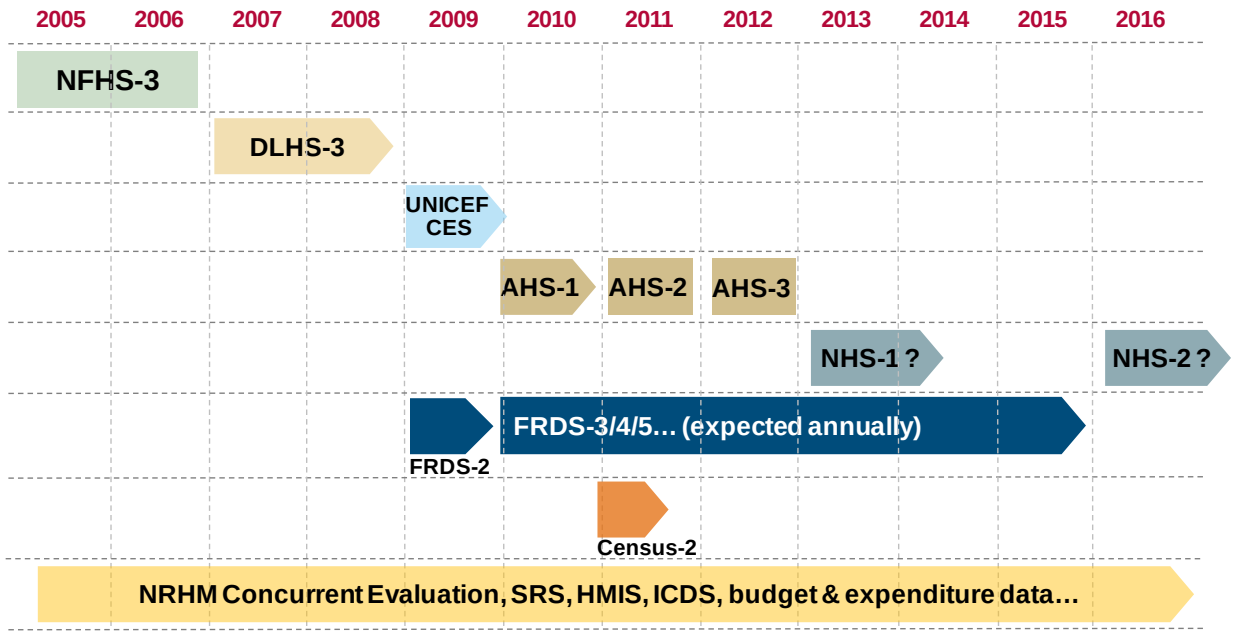
Figure IV.1 provides an overview of potential sources of relevant household survey data for the impact analysis. As shown in the figure, the most recent rounds of data collection for the NFHS and DLHS were conducted in 2005–2006 and 2007–2008, respectively. Although these data contain several relevant measures across multiple family health domains, they are relatively dated—particularly given rapid improvements in health outcomes in Bihar over the past few years—and there are no current plans to field these surveys in Bihar in the future. The NFHS, UNICEF's Coverage Evaluation Survey 2009 (CES-2009), and/or DLHS data might be able to provide additional pre-intervention data points for trend analyses, though use of multiple data sources for trend analyses can be problematic due to differences across surveys in the sample design, measures, and other factors.

A potentially more promising data source for the impact analysis was the AHS. The office of the Registrar General of India (RGI) conducted the first AHS survey in Bihar in 2010; it is anticipated that district-level descriptive statistics based on these data will be released during the second half of 2011. The second round of AHS data will be collected in 2011 and the last round is planned for 2012. The first round of AHS data just being released covers the period 2007-2009, and the third round will cover the period 2009-2011. As a result, even the last round of AHS data planned will essentially only provide "baseline" estimates for grantees, as it picks up information through 2011, and the grantees are just starting implementation in the end of 2011 and early 2012. Therefore, the AHS will not be able to provide midline data for the impact analysis as there are no current plans to collect AHS data after 2012. , precluding the use of AHS data to measure any impacts. In addition to the uncertainty around post-2012 rounds of AHS data, there are two other key reasons why the AHS will not be sufficient for the impact analysis. First, the RGI's office has stated that they cannot release the household-level AHS data files for research purposes.[26] As a

[26] The foundation has asked the GoB to request access to the household-level AHS files from RGI's office for the evaluation of the Ananya program. However, the policy of the RGI is to not provide public access to household- or individual-level data.

result, any impact analysis based on AHS data would be limited to an analysis of the district-level descriptive statistics published by the RGI after each round of data collection. Second, the AHS data also do not capture the more proximal indicators that are of interest for this evaluation, or data on family–frontline worker interactions.

Figure IV.1. Potential Secondary Data Sources for the Impact Evaluation



NFHS = National Family Health Survey; DLHS = District-Level Household Survey; UNICEF CES = UNICEF Coverage Evaluation Survey ; AHS = Annual Health Survey; NHS = National Health Survey; FRDS = Formative Research and Development Services—Routine Immunization Survey of Bihar; SRS = Sample Registration System; HMIS = Health Management Information System; ICDS = Integrated Child Development Services.

Other potential sources of survey data for the impact analysis include the Routine Immunization Survey of Bihar, collected by Formative Research and Development Services, and the Concurrent Evaluation of the National Rural Heath Mission. Although these surveys contain a more limited set of relevant family health measures than the NFHS, DLHS, and AHS, we will also explore the possibility of accessing these data, which could be valuable for supplementary analyses. It is likely that the Sample Registration System (SRS) in the future may try to collect district level estimates on key impact indicators such as IMR and under 5 mortality, and may be used to assess program impacts on these outcomes.

B. Primary Data Collection

Due to the lack of availability of existing data sources, we believe that it is essential to collect primary baseline, midline, and endline data for the impact analysis. The broad scope of the Ananya program—which is designed to improve a range of demand-side, supply-side, and health outcomes across multiple family health domains—necessitates a similarly broad-based data collection effort. The primary data sources for the impact analysis will include household, frontline worker, and facility/provider surveys. To the extent possible, we will use questions from existing, validated surveys conducted in India to develop the instruments for each of these surveys. As mentioned in Chapter III, we expect to field these surveys in fall 2011 (to establish a statewide baseline), fall 2013 (midline data to assess early impacts in the eight focus districts), and fall 2015 (to assess the overall

effects of the program).[27] The remainder of this section provides a brief overview of the design and content of the instruments for the household, frontline worker, and facility/provider surveys.

1. Household Surveys

The household survey instrument is designed to capture data on outcomes in the following broad areas: (1) access, use, and content of antenatal, delivery, and postpartum care; (2) preventative family health practices; (3) interactions with frontline workers; (4) knowledge and attitudes about key family health topics; and (5) exposure to and understanding of family health messages disseminated through various channels, and (6) neonatal mortality. Additional sections of the instrument will obtain data on various characteristics of the household and respondents that will be used to support a range of descriptive and multivariate analyses, including household composition and assets and a range of individual-level socioeconomic and demographic characteristics. Table IV.1 presents a list of the types of measures that the household survey will capture.

Most of the interventions being implemented by the grantees under the Ananya program will focus on the window between the last trimester of pregnancy and when the child is 1 year of age. In alignment with this intervention focus, the target population for the household survey will be women who have given birth within the past 12 months. We considered additionally targeting the broader population of women of reproductive age, but determined that concentrating our data collection efforts on the program's and grantees' primary target population would be a wiser use of available resources. We plan to use existing survey data sources, as well as qualitative data collected through the process evaluation, to assess the effects of the program on outcomes pertaining to this broader population.

Before conducting interviews for the household survey, all households in the selected SSUs will be listed. The listing operation will include visiting every household in each of the sampled SSUs and, for each female member of the household between the ages of 12 and 50, recording information on live births, still births, miscarriages and abortions within the last 12 months. Women who had a live birth (regardless of whether or not the child is alive at the time of the listing) will be selected for the detailed household survey. The listing information will also be used to generate estimates of stillbirths (though these households will not be selected for the detailed household survey).

2. Frontline Worker Surveys

Surveys of frontline workers, including ASHAs and AWWs, will focus on capturing outcomes data in the following domains: (1) perceived roles and responsibilities; (2) training, motivation and capabilities; (3) frequency, quality, and content of service provision; and (4) knowledge and attitudes about key family health topics. In addition, we will collect extensive data on the socioeconomic and demographic characteristics of frontline workers, including age, religion, caste, qualifications, and tenure in current post, among others (Table IV.2). The frontline worker survey will also provide valuable insight into barriers to and use of existing family health services at the community level.

[27] As noted in Chapter III, for the 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. We expect these special studies will include a one-year follow-up to provide timely information to inform the decision of whether to scale up those solutions.

Table IV.1. Household Survey: Data Elements

Household Roster
- Names of household members
- Relationship to household head
- Sex
- Age
- Marital status
- Pregnant or given birth in past year

Demographics
- Time to walk to village center
- Closest health facility/Time to reach
- Ever attended school/Highest standard completed
- Religion/Caste

Household Assets
- Material of floor/walls/roof
- Assets owned
- Source of lighting/water/cooking fuel
- Type of toilet
- Number of rooms for sleeping
- Bank account/Post office account
- Land and livestock ownership
- Food security

Pregnancy and Antenatal Care (ANC)
- Ever pregnant/Currently pregnant
- Received ANC during this/most recent pregnancy
- If not, why not
- Timing of first ANC
- Who provided/Where received ANC
- ANC services/Counseling received
- Complications during pregnancy
- Benefits received from anganwadi center
- Birth preparedness

Birth History
- Ever given birth
- Name of each baby born
- Was the birth single/multiple
- Date of birth
- Child still living
- If dead, date of death

Delivery

For Most Recent Live Birth:
- Where did delivery occur

For Institutional Delivery:
- When went to facility (stage of labor)
- Who decided to go to facility
- Mode of transportation
- Length of stay in facility
- Financial assistance received through Janani Suraksha Yojana scheme

For Home Delivery:
- Reason did not go to facility

For All Respondents:
- Who assisted delivery
- Type of delivery
- Who was present during delivery
- Safe/Clean delivery procedures used
- Complications during delivery
- Cord care
- Baby weight
- Thermal protection procedures
- Skin-to-skin contact

Post-Partum and Well-Baby Care
- Any post-partum check-up
- Timing of first check-up
- Who performed first check-up
- Where check-up performed
- Content of check-up
- Timing of first well-baby check-up
- Who performed first check-up
- Where check-up performed
- Content of check-up
- Ever breastfed/When started
- Exclusive breastfeeding
- Complementary feeding
- Immunization
- Vitamin A supplementation

Family Planning
- Current/Most recent pregnancy wanted/unwanted
- Desire for another child
- Methods known/used/source
- BBC campaign exposure
- Frontline worker ever talked about family planning

Frontline Worker Experience
- Ever contacted by frontline worker
- How often visited
- Where visited
- Information received from frontline worker
- Quality of frontline worker interaction
- Services performed by frontline worker

Facility Experience
- Ever visited facility for antenatal, delivery, post-partum, or newborn care, or family planning
- Services received in facility
- Perceived quality of facility environment (cleanliness, timeliness)
- Health personnel visited
- Perceived quality of interactions with personnel

Knowledge
- Danger signs during pregnancy, delivery, post-partum/source of information
- Birth preparedness/Source of information
- Heard about BBC campaign
- Response to BBC campaign
- Emergency preparedness/Source of information
- Clean cord care/Source of information
- Thermal care/Source of information
- Immediate breastfeeding/Source of information
- Preventative post-natal care/Source of information
- Exclusive breastfeeding/Source of information
- Complementary feeding/Source of information
- Child immunization/Source of information

Table IV.2. Frontline Worker Survey: Data Elements

Frontline Worker Background Characteristics
- Place of residence
- Sex
- Age
- Marital status
- Main occupation
- Qualifications
- Village population/Coverage
- Ever attended school/Highest standard completed
- Religion/Caste
- Health/Communications/Transportation services in village

Roles, Responsibilities and Service Provision
- Services provided
- Client profile
- Time use
- Outputs/Services provided in past 30 days
- Information/Counseling provided to clients

Training, Supervision, and Supplies
- Selection process
- Initial training
- Continuing training
- Supervision
- Equipment and supplies
- Incentives/Payment

Knowledge of Intervention Topics and Media Exposure
- Danger signs during pregnancy, delivery, post-partum/Source of information
- Facilities for antenatal care (ANC), delivery services
- Birth preparedness/Source of information
- Emergency preparedness/Source of information
- Clean cord care/Source of information
- Thermal care/Source of information
- Immediate breastfeeding/Source of information
- Preventative post-natal care/Source of information
- Exclusive breastfeeding/Source of information
- Complementary feeding/Source of information
- Child immunization/Source of information

Attitudes/Beliefs About Maternal and Child Health
- Best age for women's first pregnancy
- Opinion on content and location of ANC
- Opinion on location of delivery
- Opinion on content of preparations to ensure safe delivery
- Opinion on content of post-delivery care
- Opinion on cord care
- Opinion on skin-to-skin care
- Opinion on breastfeeding and complementary feeding
- Opinion on family planning
- Opinion on immunizations
- Opinion on sources that would be most appropriate for delivering messages on reproductive, maternal, and neonatal health

3. Facility Surveys

The objective of the facility/provider survey is to capture data that can be used to measure the effects of grantees' facility-based activities, as well as to serve as control or explanatory variables in multivariate analyses. The facility assessment component of the facility survey will collect data on (1) facility infrastructure; (2) staffing; (3) management support systems (including data collection and use and quality assurance procedures); (4) patient communication and education; and (5) the provision of antenatal, delivery, postpartum/newborn, and family planning services, including the availability of equipment, supplies, and drugs and infection-control procedures. The provider survey component will focus on training, knowledge, and attitudes, and the content and quality of care processes related to pregnancy- and newborn-related care and postpartum family planning. Table IV.3 provides a preliminary list of data elements to be included in the facility survey. This list will be refined after IFHI, led by CARE, finalizes its facility-level interventions, which were under discussion at the time of this report.

Table IV.3. Facility Survey: Data Elements

General Information
- District
- Block
- Rural/Urban
- Distance to community health center and district hospital
- Number of subcenters served
- Number of villages served
- Size of catchment area
- Hours of operation
- Trained health provider present at all times
- Routinely admits inpatients for treatment
- Number of beds
- Rooms
- Beds for overnight observation
- Overall condition of building
- Overall cleanliness of rooms and wards

Facility Infrastructure
- Electricity availability
- Water source and availability
- Functioning toilet or latrine for clients
- Communication infrastructure
- Disinfection of equipment and supplies
- Disposal of contaminated waste
- Cleanliness of facility
- Vehicle
- Laboratory facilities

Staffing
- Staff sanctioned and in position
- Number of day and night duty shifts
- Number and type of staff providing services

Management Support Systems
- Records and registries
- Number of staff meetings to discuss management/administrative issues per year
- Existence of routine program for monitoring quality of care
- Quality-of-care (QoC) procedures and protocols
- Charges for Services
- Supervision visits and procedures
- Referral procedures

Patient Communication and Education
- Availability of information, education, and communication materials on key family health topics and the method and frequency of their distribution

Service Provision

Antenatal and Postpartum Care
- Availability of ANC services
- Elements of an antenatal check-up
- Separate register for ANC patients
- Number of ANC visits per week
- Register for postpartum visits
- Number of postpartum visits in past 12 months
- Environment for ANC and postpartum exams
- Equipment for ANC and postpartum exams
- Existence of protocols/educational materials
- Mode of transportation to referral facility during obstetric emergencies
- Mode of transportation to another facility during obstetric emergencies

Normal Delivery and Newborn Care
- Availability of delivery services
- Staffing for delivery services
- Register for delivery services
- Number of deliveries in past 12 months
- Percentage of deliveries in catchment area conducted by facility
- Do auxiliary nurse midwives (ANMs) routinely provide home deliveries or attend home delivery emergencies as part of the facility service?
- Existence and content of home delivery bag
- Delivery room environment
- Equipment and supplies for deliveries
- Equipment and supplies for immediate newborn care
- Drugs stocked in the delivery room
- Protocols/Educational materials related to delivery and newborn care
- Observations made and recorded as labor is monitored
- Number of staff given labor room/newborn corner/maternity ward duty
- Rooming-in practices for this facility
- Routine administration of Vitamin A before discharge
- Routine practices for newborn care

Complicated Deliveries
- Ability to handle assisted deliveries using forceps or vacuum extractor
- Ability to perform vacuum aspiration for a woman with retained products of conception
- Equipment used for removing retained products of conception
- Ability to perform blood transfusions
- Facility conducts regular reviews of maternal or newborn deaths (or near-miss deaths)
- Experience with delivery and newborn complications

Family Planning
- Availability of family planning services
- Number of intrauterine device (IUD) insertions in past 12 months
- Major complications while inserting IUD in postpartum cases
- Number of female sterilization (tubal ligation) operations in past 12 months
- Number of postpartum tubal ligation (Minilap) procedures in past six months
- Major complications while doing tubal ligation in postpartum cases
- Frequency of family planning operations camps
- Provision of post-partum tubal ligation

V. MLE TIMELINE, REPORTING, AND DISSEMINATION OF FINDINGS

The MLE component of the Ananya program will produce a wealth of information about the implementation, cost, and effects of the program over the next five years. The study's findings will have to be communicated in an effective and timely manner to various stakeholders, including the GoB, foundation, grantees, community and development partners in Bihar, donors and NGOs working in the family health arena, researchers, and other members of the international health community.

In the first section of this chapter, we present the timeline for key evaluation activities and associated reporting. We then describe our dissemination approach, which will be elaborated as part of the execution of the MLE effort.

A. Timing of Key Evaluation Activities

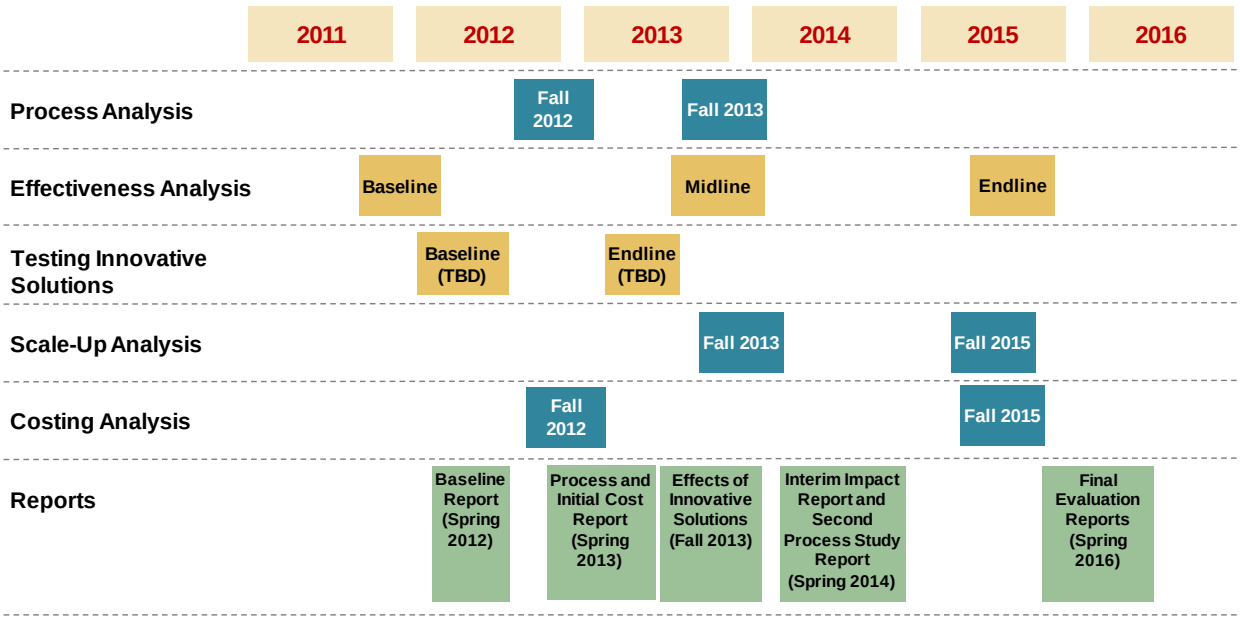
Chapter III described the timing of activities for the three main study components. Figure V.1 provides a visual summary of the timeline for major MLE activities and reporting; it does not include ongoing MLE activities that will be conducted as part of each of the study components.

- **Process and scale-up study.** We will conduct three rounds of intensive site visits and data collection as part of the process and scale-up study. The first round will be conducted in fall 2012 and will provide an early assessment of program implementation to guide program improvement and course corrections. The second and third rounds (to be conducted in fall 2013 and fall 2015, respectively) will deepen our understanding of program implementation and provide insights into how various approaches are being scaled up and sustained.

- **Impact study.** To measure the overall effects of the Anaya program, we will conduct baseline, midline, and endline surveys in the fall/winter of 2011, fall 2013, and fall 2015, respectively. After each round of data collection, we will produce a report that presents the findings of our analyses of these survey data. The baseline report will provide a benchmark against which to measure program progress; the midline report will provide information of the effects of the program in the eight focus districts; and the endline report will provide an overall assessment of the effectiveness of the Ananya program in contributing to improved family health outcomes in Bihar.

  Costing and cost-effectiveness study. We will conduct two rounds of data collection and analysis for the costing study and a cost-effectiveness analysis at the end of the program. The first round of data collection for the costing study will occur in fall 2012 and will be embedded, to the extent possible, in the data collection effort for the process analysis. These initial costing data will provide information on start-up and initial implementation costs, and will help us to refine our approach to the collection of prospective cost data. The second round of costing data collection will occur in late fall 2015 and will provide information on implementation, replication, and scale-up costs in the eight focus districts and across the state. Cost and cost-effectiveness analysis will also be conducted for innovative pilots that are rigorously tested.

Figure V.1. Timeline of Key Data Collection Activities and Reports



B. Disseminating the Findings

We will create a dissemination plan to ensure that MLE results are used for program improvement and to meet the information needs of the foundation, GoB, and other stakeholders. We will document and synthesize MLE results in ways that are relevant and accessible to key stakeholders, who will use these results for differing purposes. Information produced through the MLE effort will help grantees to refine their implementation plans to increase the effectiveness, efficiency, and sustainability of their approaches. MLE findings will also inform replication and scale-up decisions by GoB and other development partners in Bihar. Finally, evidence generated through the MLE effort can be used more broadly to inform programmatic and funding decisions by donors, NGOs, and other actors in the international health arena; to advocate for adoption of successful family health approaches; and to further the field.

We will develop and implement a communications and dissemination plan for the project life cycle that reaches and engages key stakeholders, disseminates results in real time, and promotes feedback from and interaction with internal and external audiences. For the communications and dissemination plan to be effective, we must clearly identify, at the outset, the key internal and external audiences for the various MLE results and products. We will create a tiered priority structure to categorize the various audiences, which will acknowledge the fact that not all audiences are interested in the same types of information or levels of (technical) detail; it will also ensure that our dissemination plan addresses the needs all of potential audiences.

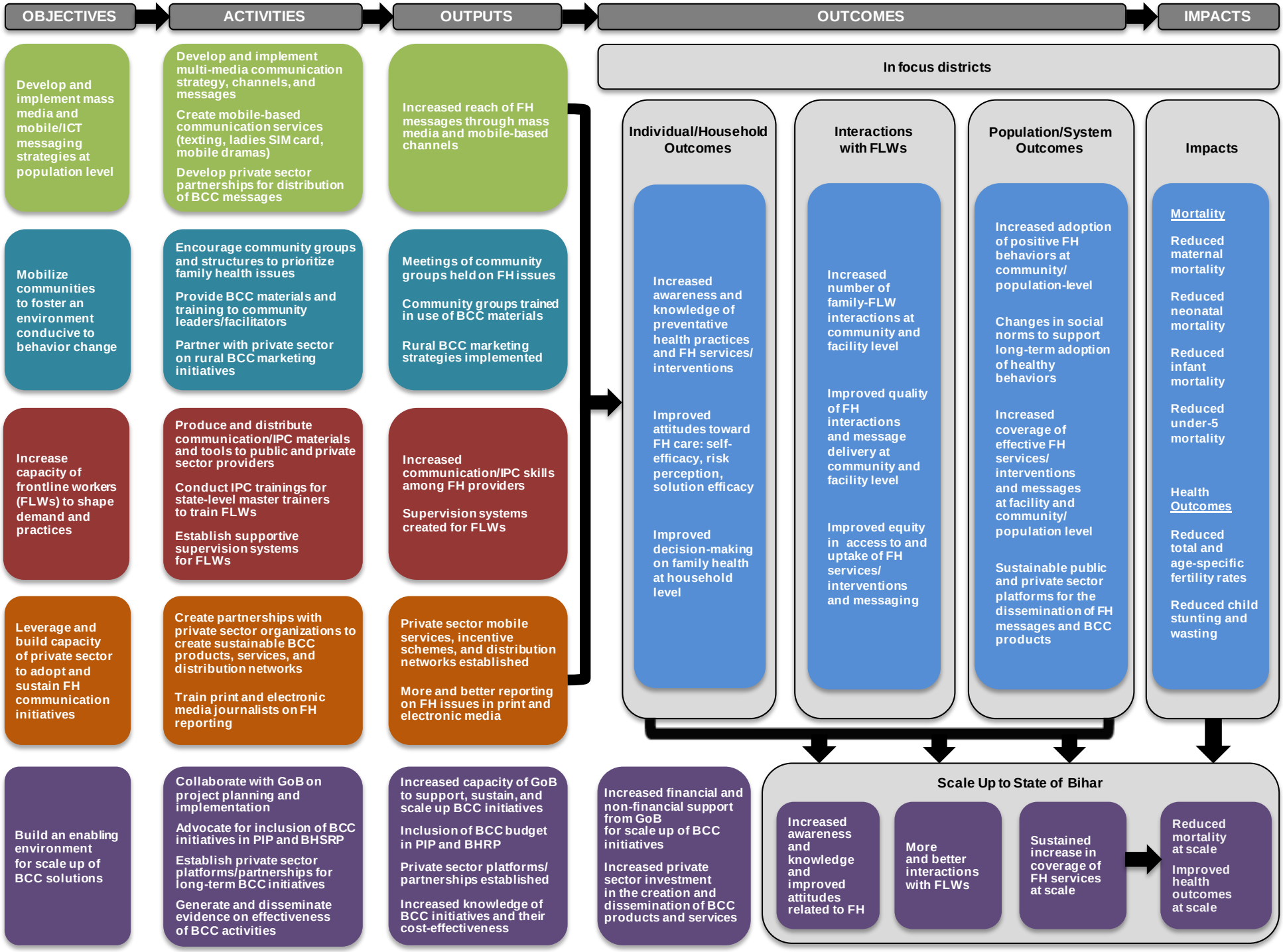
The MLE team will work with the foundation, grantees, and partners to determine the most effective and appropriate modes of communication for each audience type. For example, we will use the biannual MLE partners' workshops, which include major partners and stakeholders in Bihar, to obtain input on how best to capture and disseminate learning to inform program improvement and course correction. We anticipate producing a range of dissemination products, including reports, topic papers, research/evaluation briefs, journal articles, and targeted emails; we will also present findings at various meetings and conferences.

APPENDIX A

LOGIC MODELS FOR THE GRANTS

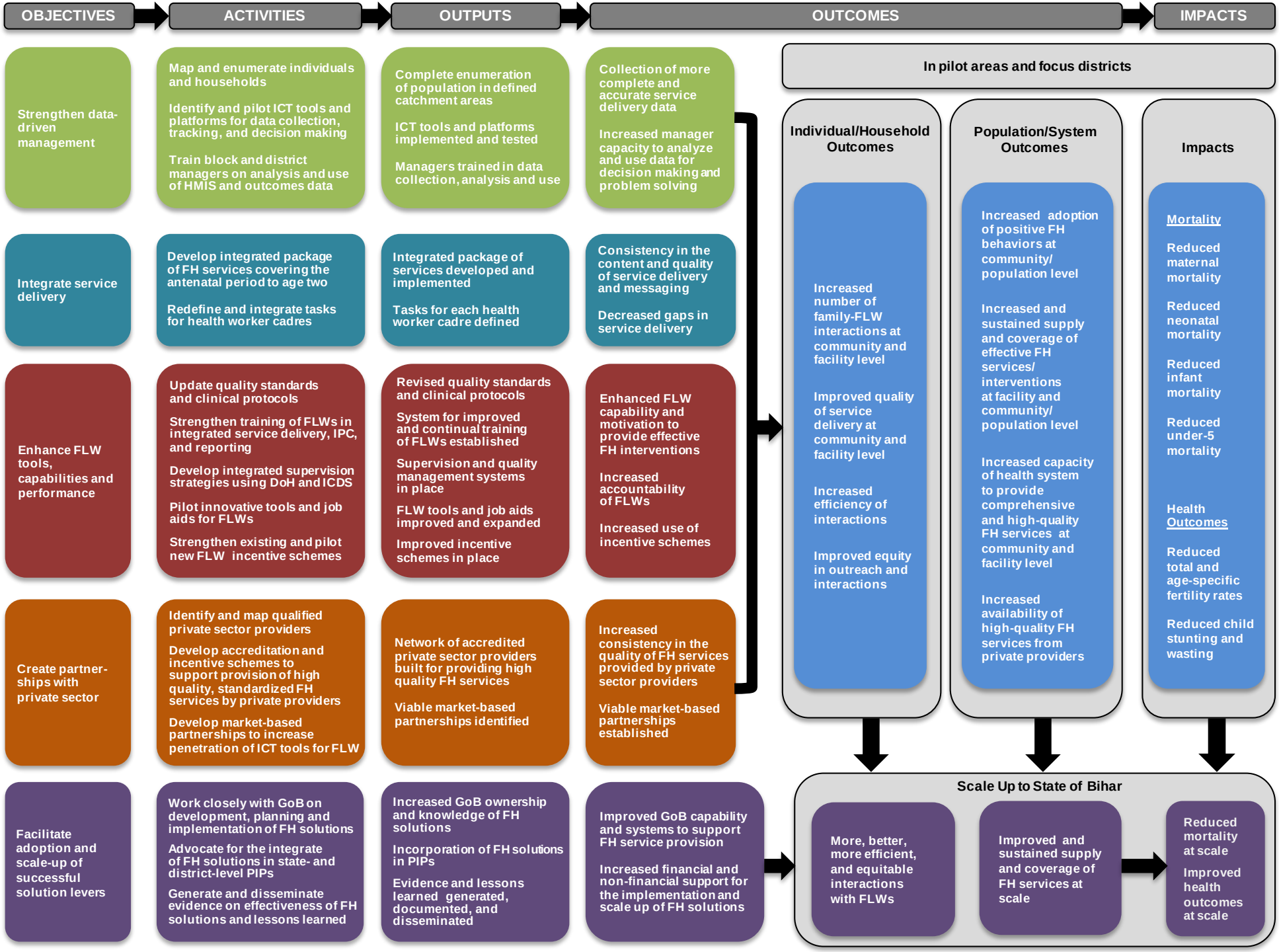
This page has been left blank for double-sided copying.

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



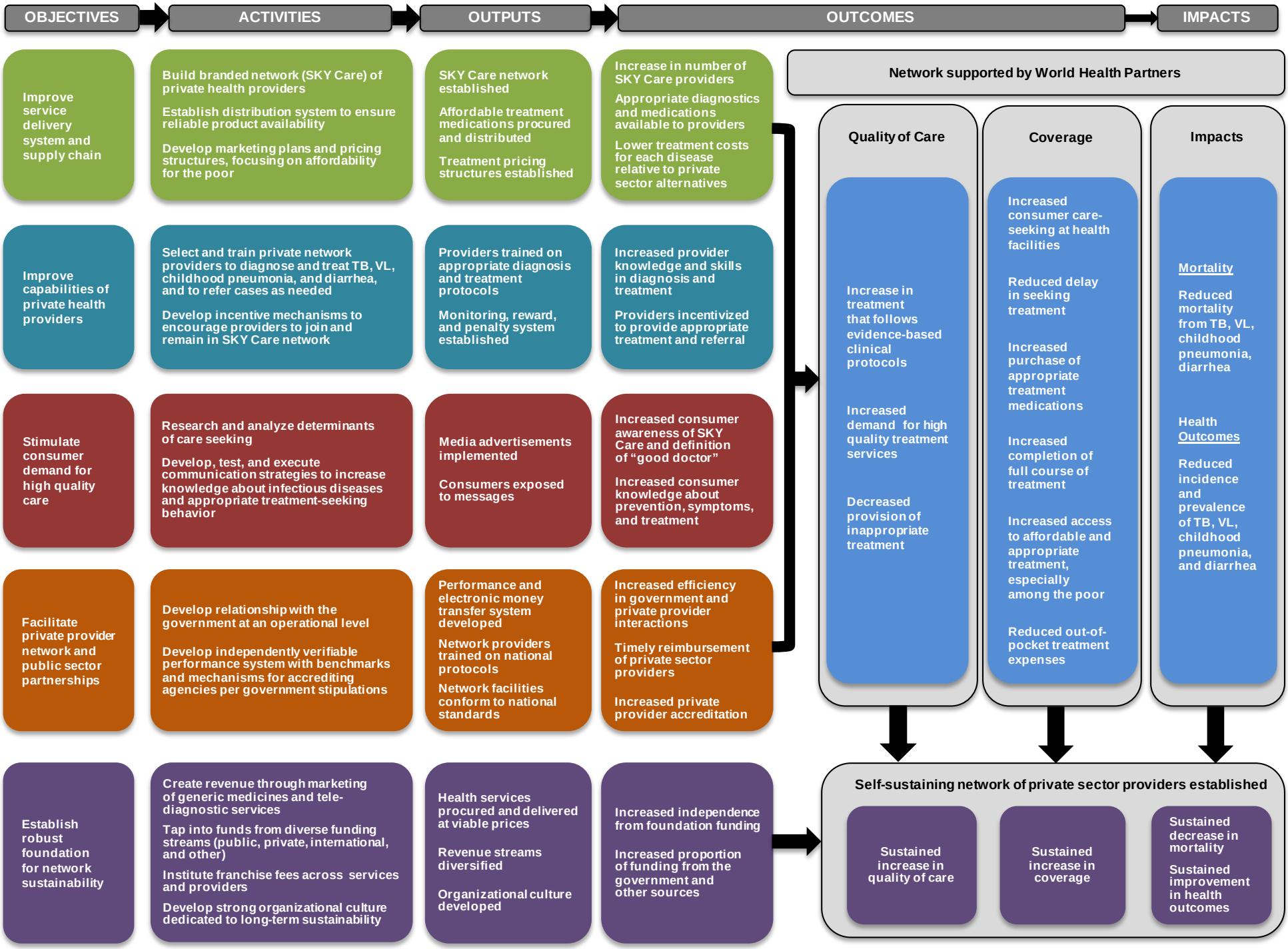
This page has been left blank for double-sided copying.

Figure A.2. Integrated Family Health Initiative in Bihar (CARE)



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

Figure A.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.

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

MATHEMATICA
Policy Research

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