# Evaluation of the Team-Based Goals and Performance Based Incentives (TBGI) Intervention in Bihar

Final Report

February 14, 2014

Evan Borkum
Anu Rangarajan
Dana Rotz
Swetha Sridharan
Sukhmani Sethi
Mercy Manoranjini

MATHEMATICA
CENTER FOR
INTERNATIONAL POLICY
RESEARCH AND EVALUATION

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

Contract Number:
4993/22539 (18)

Mathematica Reference Number:
40195.540

Submitted to:
Bill & Melinda Gates Foundation
3rd Fl., Left Wing
Capital Court Building, Olof Palme Marag
Munirka, New Delhi, India 110 067
Project Officer:
Katherine Hay, Senior Program Officer
Usha Kiran Tarigopula, Senior Program Officer

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

# Evaluation of the Team-Based Goals and Performance Based Incentives (TBGI) Innovation in Bihar

Final Report

February 14, 2014

Evan Borkum
Anu Rangarajan
Dana Rotz
Swetha Sridharan
Sukhmani Sethi
Mercy Manoranjini

MATHEMATICA
CENTER FOR
INTERNATIONAL POLICY
RESEARCH AND EVALUATION

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

## ACKNOWLEDGMENTS

We are indebted to a broad range of individuals and organizations for their support of this effort. At the Bill & Melinda Gates Foundation, we would like to thank Katherine Hay, Usha Tarigopula, and Yamini Atmavilas for their helpful feedback on drafts of the report as well as their guidance throughout the project. We would also like to express our gratitude to the team at CARE, who developed and implemented this innovative intervention and without whom this evaluation would not have been possible. Shamik Trehan, Indrajit Chaudhuri, Praveen Bhalla, Benjamin Schwartz, Sunil Mohanty and his district team, and colleagues at the Georgia Institute of Technology provided valuable input on our questionnaires and this report, and spent many hours meeting with our researchers in the field to discuss the intervention and share their experiences.

We are grateful, as always, to the team at Sambodhi Research and Communications for leading a strong data collection effort. Kultar Singh, Aparna Seth, Chandan Singh, Nikhil Raj, and Srishti Dayal diligently piloted and revised multiple rounds of our survey instruments, provided an in-depth training to field investigators, and surmounted considerable logistical challenges—chief among them being heavy monsoons and extensive flooding—to successfully administer the surveys. The contributions of several Mathematica staff were also instrumental in this effort. Our programmers Carol Razafindrakoto, Seth Morgan, Bethany Simard, and Connie Qian worked hard to clean and analyze the data. Nancy Murray reviewed the draft report and provided thoughtful comments. John Kennedy expertly edited the report, and Jennifer Baskwell assisted with formatting and production.

Finally, we would like to thank the households and frontline workers who took part in this study. Their willingness to answer our detailed questions and share their opinions and experiences are what finally made this study possible.

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

## CONTENTS

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

# TABLES

This page has been left blank for double-sided copying

.

# FIGURES

# EXECUTIVE SUMMARY

The Team-Based Goals and Performance-Based Incentives (TBGI) intervention, which CARE conceptualized, developed, and implemented as part of the Ananya program in Bihar, leverages the power of incentives and lessons from motivational theory on teamwork and goal-setting to help improve maternal and child health. Under the intervention, CARE set targets for the percentage of eligible beneficiaries in a subcenter catchment area who should have adopted each of seven key health behaviors or goals (Box 1). All frontline health workers (FLWs) in a given subcenter, including the accredited social health activists (ASHAs), Anganwadi workers (AWWs), and the subcenters' auxiliary nurse midwives (ANMs), received nonmonetary incentives (consisting of small household items) if their subcenters met five of seven goals in a given quarter. The intervention explicitly sought to encourage teamwork and cooperation among FLWs by providing these incentives for achievements by the subcenter as a whole rather than by individual FLWs, and by providing FLWs with information on the concept and importance of teamwork. It included additional elements to motivate the FLWs in each subcenter, such as a service pledge they recited together and a certificate of recognition for subcenters that met their targets in all quarters. Overall, the intervention was expected to lead to improvements in the incentivized outcomes and to broader changes in related, but nonincentivized, outcomes through increased FLW motivation and teamwork.

**Box 1. The Seven TBGI Goals Include:**

1. Pregnant women who had made arrangements for transportation for their delivery
2. Pregnant women who received at least 90 iron/folic acid (IFA) tablets in their final trimester of pregnancy
3. Children who were breastfed within an hour of their delivery
4. Deliveries in which appropriate umbilical cord care procedures were followed
5. Children in the age group of 6 to 11 months who are fed food that is age-appropriate and nutritious
6. Women (or their partners) who used any modern method of family planning within 6 months of delivery
7. Children who received a DPT3 injection within 6 months of birth

We conducted a rigorous evaluation of the impacts of the TBGI intervention through a clustered randomized controlled trial (RCT) in Begusarai district of Bihar. The pilot intervention began in August 2012. The RCT involved randomly assigning 76 subcenters to a treatment group of 38 treatment subcenters that received the intervention and a control group of 38 subcenters that did not. The RCT enabled us to measure the value-added of TBGI beyond the other Ananya program interventions, which were implemented simultaneously in both the treatment and control groups. Specifically, the evaluation sought to address the following research questions:

- What was the TBGI intervention and how was it implemented?
- What was the impact of the TBGI intervention on FLW–household interaction?
- Did the intervention lead to improvements in maternal and child health outcomes among beneficiaries?

To address these questions, we used a mixed-methods approach, which collected and analyzed quantitative and qualitative data. The quantitative approach used an RCT design to examine whether the intervention led to changes in how FLWs provided services and in household behaviors and practices, based on data from surveys conducted with about 640 FLWs and 1,600 beneficiaries approximately one year after the introduction of the intervention. The qualitative analysis examined questions related to the implementation of the program based on field visits and semistructured interviews with implementing partner staff, FLWs, and beneficiaries. Next, we summarize our key findings from the evaluation.

**FLWs generally understood the nature of the intervention, but CARE played an important role in determining targets and monitoring progress.**

Our interviews and discussions with the FLWs suggested that most understood the nature of the intervention, including the setting of targets for specific outcomes and the number of targets they had to meet to receive incentives. However, in both the qualitative interviews and FLW surveys, FLWs articulated a varied understanding of what accomplishing each of the specific seven goals meant. For example, 35 percent of FLWs felt they achieved the IFA goal if they gave a woman the tablets, but 70 percent thought it was necessary to ask the woman if she had consumed the tablets. In addition, although most FLWs knew the numerical targets needed to reach the percentage targets, about one in three required assistance from CARE staff with these calculations.

Not unexpectedly, given that CARE implemented the intervention, its staff also played an important role in helping FLWs to assess progress toward targets. For example, 73 percent of AHSAs and AWWs in treatment subcenters reported that someone from CARE reviewed their TBGI diaries—which AWWs used to track progress—in the past month. In addition, ANMs in treatment subcenters were more likely than those in control subcenters to report having received assistance from CARE staff in subcenter meetings, which the main Ananya program introduced but also served as a forum to monitor and discuss progress toward goals in the treatment subcenters under TBGI. Specifically, ANMs in treatment areas were more likely to report that they received assistance from CARE staff in reviewing FLW registers during the most recent subcenter meeting (13 percent of treatment group ANMs reviewed registers on their own, compared with 31 percent of the control group, a statistically significant difference) and in leading the most recent meeting (20 percent of treatment group ANMs led the most recent meeting, compared with 35 percent in the control group, although this difference was not significant).

**The TBGI intervention affected attendance at and focus of subcenter meetings, and increased teamwork and coordination among FLWs.**

The TBGI intervention led to increased attendance at subcenter meetings (Table 1). Although 82 percent of control ASHAs and AWWs reported attending all monthly submeetings in the three months before our surveys, 94 percent of treatment ASHAs and AWWs reported doing so, a statistically significant difference. ASHAs and AWWs in treatment subcenters were also significantly more likely to have discussed certain topics directly linked to TBGI targets in the most recent subcenter meeting compared with the control group, including IFA tablets (90 percent in treatment compared with 83 percent in control), umbilical cord care (69 percent in treatment compared with

53 percent in control), and immediate breastfeeding (78 percent in treatment compared with 62 percent in control).

The TBGI intervention also influenced how FLWs in a subcenter worked together (Table 1). FLWs in the treatment subcenters were significantly more likely to report that they were expected to always meet and plan with their teams relative to the control group. ASHAs and AWWs in treatment subcenters also met significantly more often with one another to discuss work compared with their counterparts in the control subcenters. In addition, ASHAs and AWWs in treatment subcenters were more likely than their control counterparts to view the ANM at their subcenter as part of their team (87 percent in treatment compared with 79 percent in control), and reported that the ANM was more often available to assist them and more commonly gave helpful advice on how to handle difficult cases.

The increase in teamwork and coordination was evident in the conduct of home visits by FLWs (Table 1). Specifically, ASHAs and AWWs in treatment subcenters reported conducting significantly more joint visits than their counterparts in control subcenters. These findings are consistent with beneficiary reports, which suggested that treatment beneficiaries were more likely to report joint visits by FLWs (as described later). They are also consistent with the qualitative data, which suggest that the TBGI intervention led to increased teamwork between ASHAs and AWWs—including through joint visits—and increased willingness for ASHAs and AWWs to reach out to ANMs for assistance as needed.

**Table 1. Impacts of TBGI on ASHAs and AWWs**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **Subcenter Meetings** | | | | |
| ASHA/AWW Attended Three or More Subcenter Meetings in Past Three Months | 82.7 | 93.9 | 11.2** | 0.000 |
| Topics Discussed at Subcenter Meetings | | | | |
| IFA tablets | 83.2 | 89.7 | 6.5** | 0.032 |
| Arranging transportation | 99.1 | 99.6 | 0.5 | 0.569 |
| Cord care | 52.9 | 69.4 | 16.4** | 0.000 |
| Immediate breastfeeding | 61.8 | 77.5 | 15.7** | 0.000 |
| Feeding child semisolid food | 98.6 | 100.0 | 1.4 | 0.129 |
| Family planning | 97.9 | 100.0 | 2.1* | 0.054 |
| **Working in a Team** | | | | |
| Always Expected to Plan with Team[a] | 52.5 | 68.8 | 16.3** | 0.002 |
| Always Expected to Meet Regularly with Team[a] | 56.1 | 65.3 | 9.2** | 0.049 |
| Average Times Met with ASHA/AWW (opposite cadre) in Past Week to Discuss Work | 1.53 | 2.14 | 0.61** | 0.000 |
| **Working with Subcenter ANM** | | | | |
| Considers ANM as Part of the Subcenter Team | 78.9 | 86.7 | 7.8** | 0.038 |
| ANM Available When Needed Most of the Time[b] | 58.7 | 67.9 | 9.2* | 0.068 |
| ANM Gives Helpful Advice on Certain Cases Most of the Time[b] | 41.3 | 52.8 | 11.5** | 0.031 |
| **Working with ASHA/AWW (opposite cadre)** | | | | |
| Ever Conduct Joint Visit | 69.0 | 78.3 | 9.3** | 0.017 |
| Average Number of Joint Home Visits in Past Week | 1.24 | 1.77 | 0.53** | 0.000 |
| **Sample Size** | **273** | **273** | | |

Source: Ananya TBGI RCT endline survey.

Notes: Item-specific nonresponse might limit sample size for some comparisons. Adjusted difference is from a regression analysis using a linear probability model with controls for subcenter and FLW characteristics (see Chapter IV for details).

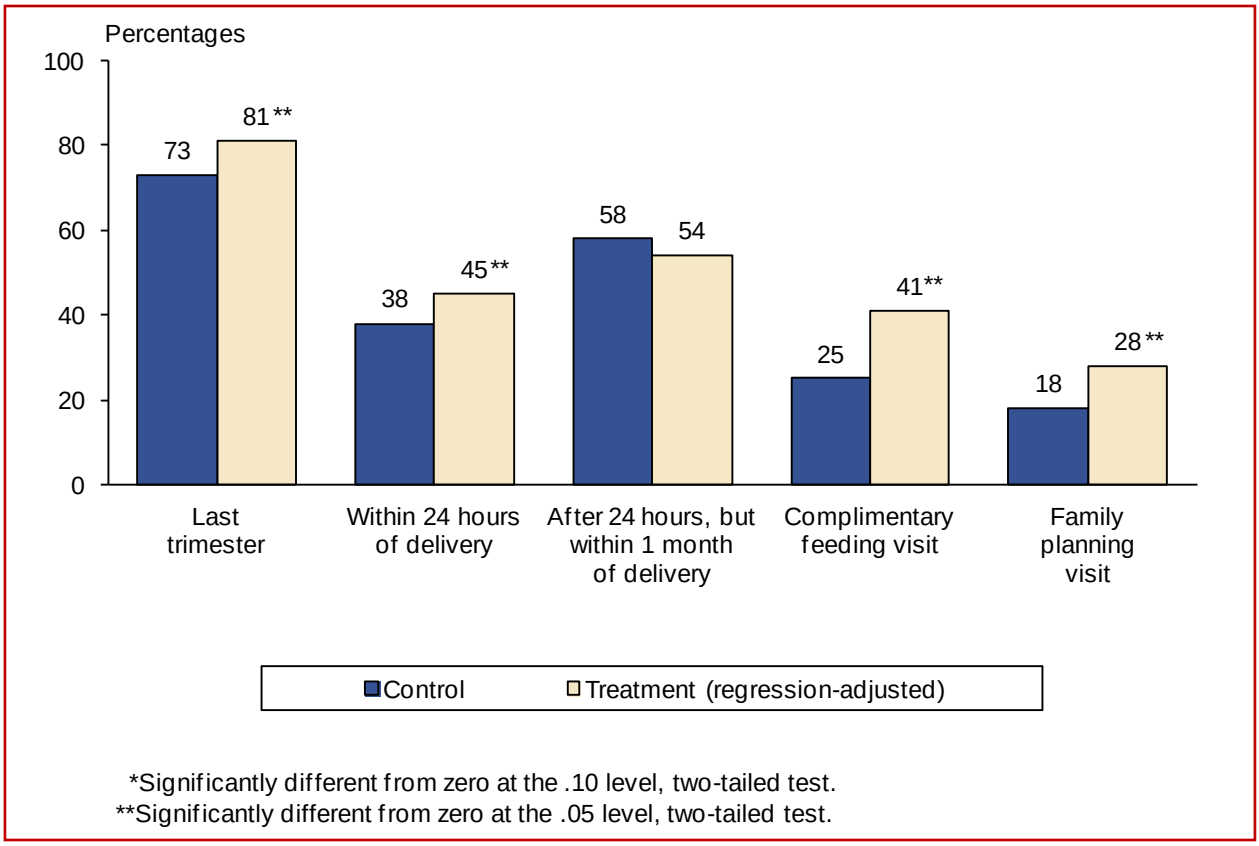
[a] Response options are always, sometimes, and never.
[b] Response options are most of the time, some of the time, and none of the time.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

**The TBGI intervention improved the frequency of FLW–beneficiary interactions and the advice provided during home visits.**

The percentage of beneficiaries reporting they received any visit from an FLW at a critical time during pregnancy and early childhood was significantly higher in the treatment group relative to the control group (Figure 1). In the absence of the intervention, 73 percent of the control group reported an FLW home visit during the final trimester of pregnancy and 38 percent reported a visit in the first 24 hours after delivery; in the treatment group, these rates were both about 7 percentage points higher (about 10 and 18 percent of the respective control means). Impacts on FLW visits around complementary feeding for children 5 months or older were uniformly large, positive, and strongly significant. For example, 25 percent of beneficiaries in the control group received such a visit compared with about 41 percent of beneficiaries in the treatment group (a significant difference of 16 percentage points, or 64 percent of the control mean). Similarly, beneficiaries in the treatment group were far more likely to have received advice on issues such as the initiation of complementary feeding at age 6 months, the type of food, and the frequency and quantity of feeding—all areas emphasized by the TBGI goal around child nutrition. In addition, 28 percent of the treatment group reported a visit by an FLW to discuss family planning compared with 18 percent of the control group (a significant difference of 10 percentage points, or 55 percent of the control mean), which relates to the TBGI goal promoting use of modern contraceptive methods. These findings are consistent with the qualitative findings that FLWs used the TBGI goals to focus the messages that they delivered during home visits.

**Figure 1. Home Visits by FLWs**



Consistent with the results from the FLW surveys and qualitative data, beneficiaries in the treatment group more commonly reported joint visits by FLWs. About 37 percent of treatment beneficiaries reported having received a joint home visit from an AWW and ASHA compared with 27 percent of control beneficiaries, whereas 32 percent of treatment beneficiaries reported having received a joint visit from the ANM and another FLW compared with 23 percent of control beneficiaries; both of these differences are statistically significant and large in magnitude (almost 40 percent of the control means).

**Impacts on beneficiaries' behaviors were, however, much more modest, but some evidence suggests positive impacts for certain outcomes.**

The TBGI intervention directly incentivized FLWs based on their performance around the seven TBGI goals. Although we see large differences in FLW–beneficiary interactions and discussions around these behaviors, the differences between the treatment and control groups in the outcomes most directly reflecting these goals were generally much more modest, below the magnitude of 5 to 10 percentage points that the evaluation was powered to detect (Table 2). Nevertheless, there was some evidence of impacts for certain incentivized outcomes as well as for other nonincentivized outcomes. First, there were statistically significant improvements in some specific feeding practices related to the goal around complementary feeding. For example, 55 percent of children in the treatment group were fed a cereal-based meal in the previous day versus 46 percent of the control group; 41 percent of the treatment group were fed from a separate bowl versus 32 percent of the control group. Second, although the impact on DPT3 was not statistically significant, it was relatively large in magnitude (6 percentage points). Third, there was a positive and significant impact on the use of modern contraception methods for women who gave birth *more* than six months ago (26 percent in the treatment group versus 15 percent in the control group, a difference of 11 percentage points or 73 percent of the control mean), although the specific TBGI goal focused on the use of modern contraception for mothers within six months of giving birth. Fourth, we also found some evidence of a significant and positive impact on exclusive breastfeeding, a nonincentivized outcome. Overall, these results suggest that the TBGI has the potential to add value to the existing Ananya interventions and lead to larger changes in household behaviors at relatively low additional cost (given the low cost of the gifts provided as incentives), though translating improved FLW–household interactions from TBGI into major behavior change is an ongoing challenge.

**Further strengthening of ANM leadership skills and greater clarity to FLWs about the meaning of attaining a goal will be important if, or when, the program scales up.**

Our evaluation, which tracked outcomes over a one-year period after the start of the intervention, shows large improvements in FLW–beneficiary interactions as a result of this intervention, but more modest changes in beneficiaries' behaviors. It is possible that, over time, some of these beneficiaries' behaviors might improve as they hear FLWs' messages more and more, and as the FLWs improve their ability to communicate these messages. On the other hand, it is also possible that the initial buzz around these incentives and teamwork will dissipate over time, so we might see lower impacts in the future. If continued, it will be important for the intervention to identify ways to maintain FLWs' enthusiasm and motivation as they get together during the subcenter meetings.

Our evaluation also provides some inputs for sustainability if the program is scaled up. First, it will be important to further train and strengthen the role of ANMs so they can lead the subcenter meetings, including communication of content to more technical information of how to calculate targets and track them. Second, FLWs might benefit from guidance on defining what it means to

achieve goals; uncertainty about these definitions might undermine FLWs' perceptions of what they have to do to meet their targets, and their ability to effect changes in beneficiaries' behaviors regarding key maternal and child health practices. Third, it will be important to consider how to handle procurement and distribution of nonmonetary incentives at much larger scale by the government of Bihar. Finally, if scaled up across the state, a practical and feasible monitoring mechanism will have to be set up to ensure accurate tracking and reporting of FLWs' achievements.

**Table 2. Impacts on Beneficiary Outcomes (incentivized behaviors highlighted)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **Antenatal Care and Delivery** | | | | |
| Obtained number of ambulance, private vehicle, or FLW for delivery transportation | 55.4 | 54.5 | -0.8 | 0.829 |
| At least 90 IFA tablets received | 16.0 | 18.0 | 2.0 | 0.516 |
| **Newborn Care** | | | | |
| Immediate breastfeeding | 56.4 | 58.4 | 2.0 | 0.613 |
| Nothing applied to cord | 56.0 | 55.6 | -0.4 | 0.928 |
| **Infant Feeding** | | | | |
| Complementary Feeding (ages 6 to 11 months): | | | | |
| Child eats solid or semisolid food | 62.4 | 67.4 | 5.0 | 0.211 |
| Child began eating solid food by age 6 months | 25.4 | 28.7 | 3.3 | 0.427 |
| Child feeding in previous day | | | | |
| Times fed | 1.24 | 1.45 | 0.21 | 0.085* |
| Fed any meal from separate bowl | 31.7 | 40.9 | 9.2 | 0.047** |
| Fed any cereal-based meal | 46.3 | 55.3 | 9.0 | 0.016** |
| Amount fed (*katoris*) | 0.35 | 0.47 | 0.11 | 0.097* |
| Exclusive Breastfeeding for 6 Months (children 6 months or older)[a] | 26.3 | 34.2 | 8.0 | 0.145 |
| Exclusive Breastfeeding in Past 24 hours (ages younger than 6 months)[b] | 61.3 | 70.5 | 9.2 | 0.018** |
| **Family Planning** | | | | |
| Women Who Gave Birth in Previous 6 Months: | | | | |
| Use of permanent methods | 7.6 | 4.6 | -3.0 | 0.073* |
| Use of any modern method | 10.5 | 10.3 | -0.2 | 0.957 |
| Women Who Gave Birth More than 6 Months Ago: | | | | |
| Use of permanent methods | 12.8 | 15.6 | 2.9 | 0.437 |
| Use of any modern method | 15.4 | 26.1 | 10.6 | 0.015** |
| **Immunization** | | | | |
| Received DPT3 (card and self-reports) | 67.4 | 72.8 | 5.4 | 0.243 |
| Received DPT3 (card only) | 68.5 | 75.8 | 7.3 | 0.113 |
| Received DPT3 (self-reports only) | 61.3 | 65.8 | 4.5 | 0.584 |
| Received DPT1 (card and self-reports) | 88.0 | 89.5 | 1.5 | 0.571 |
| Received DPT1 (card only) | 85.6 | 93.5 | 7.9 | 0.011** |
| Received DPT1 (self-reports only) | 86.4 | 84.0 | -2.4 | 0.566 |

Source: Ananya TBGI RCT endline survey.

Note: Adjusted difference is from a regression analysis using a linear probability model with controls for subcenter and beneficiary characteristics (see Chapter V for details). Sample sizes vary by outcome due to different age restrictions, and are from 261 to 1,200.

[a] Based on self-reports of the duration of exclusive breastfeeding for children 6 months or older.
[b] Based on reports of liquids and solids fed to children younger than 6 months old in the previous 24 hours, following the recommended definition of the World Health Organization.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

# I. INTRODUCTION

Interventions providing performance-based incentives to health workers for achieving specific health-related targets are a commonly used strategy to improve public health outcomes in the developing world. Using evidence from case studies around the world, Eichler et al. (2009) argue that this type of intervention has the promise to improve health. Several studies of specific performance-based incentive interventions for health care workers in developing countries have been conducted. For example, Peabody et al. (2013) examined an intervention in the Philippines that rewarded doctors based on subjective measures of performance. The authors found substantial improvements in children's health associated with the intervention. In another study, Basinga et al. (2010) explored incentives given to health centers in Rwanda for the quality and quantity of services provided, finding significant improvements in some—but not all—domains. Miller and Babiarz (2013) provide a recent and comprehensive review of strategies using performance-based incentives to improve health in low- and middle-income countries. Many of these strategies target facility-based health workers, and most rely on financial incentives—rewarding either individual providers through increased private income or facilities as a whole through increased operating budgets.

In India, several incentive programs for health workers have been or are currently being employed to improve maternal and child health. Most of these programs use financial incentives to incentivize the provision of specific services by frontline health workers (FLWs). Examples include incentivizing FLW when they are able to facilitate women to give birth at a public health facility (the Janani Suraksha Yojana, or JSY, scheme), convincing women or their husbands to undergo sterilization, and ensuring that children receive immunizations (for a summary of financial incentives available to accredited social health activists [ASHAs], one of the cadres of FLWs, see Wang et al. 2012). To the best of our knowledge, the only formal evaluation of these schemes has focused on the JSY scheme, which includes financial rewards for both FLWs and mothers when women give birth at a public health facility rather than at home, and is one of the largest incentive-based health policies implemented in India. Lim et al. (2010) provide evidence that the scheme was associated with increases in institutional deliveries and antenatal care and reductions in perinatal and neonatal mortality (though Mazumdar et al. [2011] debated the impact on mortality). Although most of the existing incentive schemes focus on provision of specific services, Singh (2011) conducted a small-scale experimental study to examine the effect of giving incentives to Anganwadi workers (AWWs) based on the overall health of the children attending their Anganwadi Center, as measured by indicators of undernutrition. He found that these bonuses improved children's weight for age. Overall, the available research in developing countries in general and in India in particular suggests that incentives for health workers might have a strong role to play in helping improve health outcomes.

The Team-Based Goals and Performance-Based Incentives (TBGI) intervention, which was conceptualized and implemented by CARE as part of the Ananya program in Bihar, leverages both the power of incentives and lessons from motivational theory on teamwork and goal-setting (Herzberg 1987; Dieleman and Harnmeijer 2006) to help improve maternal and child health. The intervention, which this report describes in further detail, differed from existing incentive schemes in several key respects. First, the intervention integrated incentives and other motivational tools, rather than relying on incentives alone—a novel approach. Second, the intervention incentivized the achievement of an integrated set of outcomes, awarding incentives based on an assessment of overall achievement of this set of outcomes. In contrast, most existing incentive schemes in India incentivize the achievement of a single outcome. The TBGI intervention's focus on the full set of outcomes aimed to avoid the possible distorting effects on health workers' behavior of outcome-

specific incentives, which might encourage health workers to focus on specific outcomes at the expense of others. Third, the TBGI incentives were nonmonetary in nature, largely consisting of inexpensive household items such as cooking utensils and hotpots, whereas most existing incentives in India and elsewhere are monetary. Although monetary incentives were considered, it was assessed that the program would have greater chance of scale-up if the incentives were nonmonetary, particularly since financial transactions need to go through more complicated appropriation processes. In addition, CARE determined through discussions with FLWs that they actually preferred noncash incentives. And because CARE was able to secure bulk discounts for the household items provided as incentives, the value of these incentives was higher than if the FLWs had received the incentive funding in cash. Finally, existing incentive schemes usually reward individuals directly for individual performance *or* reward groups of individuals for group performance through rewards that accrue to the entire group as a whole (for example, facilities might be rewarded by increased operating budgets). In contrast, the TBGI intervention incorporates features of both these approaches: award determination is made based on the performance of the subcenter as a whole, but the actual incentives are awarded to (all) frontline workers within the subcenter, regardless of an individual met their goals or not. Thus, the frontline workers need to work in close coordination with others in the subcenter and to ensure that moral hazard or free rider issues do not arise.

## A. Research Questions and Approach

The intervention was piloted in Begusarai district, and was rigorously evaluated using a clustered randomized controlled trial (RCT) design. This report summarizes our findings from this evaluation. In particular, we address the following questions:

- **What was the TBGI intervention and how was it implemented?** Did the FLWs understand the intervention? Were they motivated by it? What were the implementation successes and challenges? Are there any lessons for scale-up?
- **What was the impact of the TBGI intervention on FLW–household interaction?** Did the intervention lead to the provision of more and better services by FLWs? Did the intervention lead to greater coordination in service delivery between the FLWs?
- **Did the intervention lead to improvements in maternal and child health outcomes among beneficiaries?** Did the impacts vary by key subgroups of beneficiaries, such as those defined by caste, socioeconomic status (SES), or birth parity?

To address these questions, we used a mixed-methods approach, which collected and analyzed quantitative and qualitative data. In particular, we used an RCT design to examine whether the intervention led to changes in how FLWs provided services and in household behaviors and practices, based on data from surveys conducted with FLWs and beneficiaries. We complemented our quantitative analyses with qualitative data obtained through field visits and semistructured interviews with implementing partner staff, FLWs, and beneficiaries to examined questions related to the implementation of the intervention.

The rest of this report is structured as follows. In the remainder of Chapter I, we describe the TBGI intervention in greater detail. Chapter II describes the research design, sample, data collection, and methodological approach (Appendix A contains more details on these). Chapter III uses both qualitative and quantitative data to describe the implementation of the interventions, and in particular from the perspective of the frontline workers who were in the treatment subcenters. Chapters IV and V summarizes impacts on outcomes from surveys conducted with FLWs and

beneficiaries, respectively. We end in Chapter VI with a brief summary and conclusion, including some key lessons and considerations if the intervention is to be scaled-up.

## B. Description of the TBGI Intervention

The TBGI intervention aims to improve health outcomes of beneficiaries by motivating FLWs in each subcenter[1] to improve service provision by (1) creating a sense of collective team responsibility, solidarity, and teamwork among the FLWs; (2) defining specific goals related to coverage of household maternal and child health practices, and setting specific subcenter-level targets to achieve in each quarter; and (3) providing a small nonmonetary incentive for all FLWs in the subcenter upon attainment of the collective subcenter goals.

The TBGI intervention focuses on attaining targets for seven key goals:

1. Pregnant women who had made arrangements for transportation for their delivery
2. Pregnant women who received at least 90 iron/folic acid (IFA) tablets in their final trimester of pregnancy
3. Children who were breastfed within an hour of their delivery
4. Deliveries in which appropriate umbilical cord care procedures were followed
5. Children in the age group of 6 to 11 months who are fed food that is age-appropriate and nutritious
6. Women (or their partners) who used any modern method of family planning within 6 months of delivery
7. Children who received a DPT3 injection within 6 months of birth

These goals were purposefully selected by CARE to include a mixture of relatively straightforward goals (for example, ensuring transportation arrangements for delivery), which could be achieved largely through FLW actions such as providing the right contact information, and much more complex goals involving behavior change (such as adoption of modern methods of contraception), which might be more challenging to achieve. The intention in selecting this diverse range of outcomes was to strike a balance between targeting key family health outcomes that required improvement, while avoiding demotivating FLWs by selecting only challenging behaviors whose prevalence were at very low levels and that were difficult to change in the social context of Bihar. In addition, it was anticipated that the focus on these seven goals would prompt broader changes in nontargeted outcomes. For example, increased home visits by FLWs to discuss appropriate umbilical cord care could have included additional advice on other nontargeted aspects of newborn care.

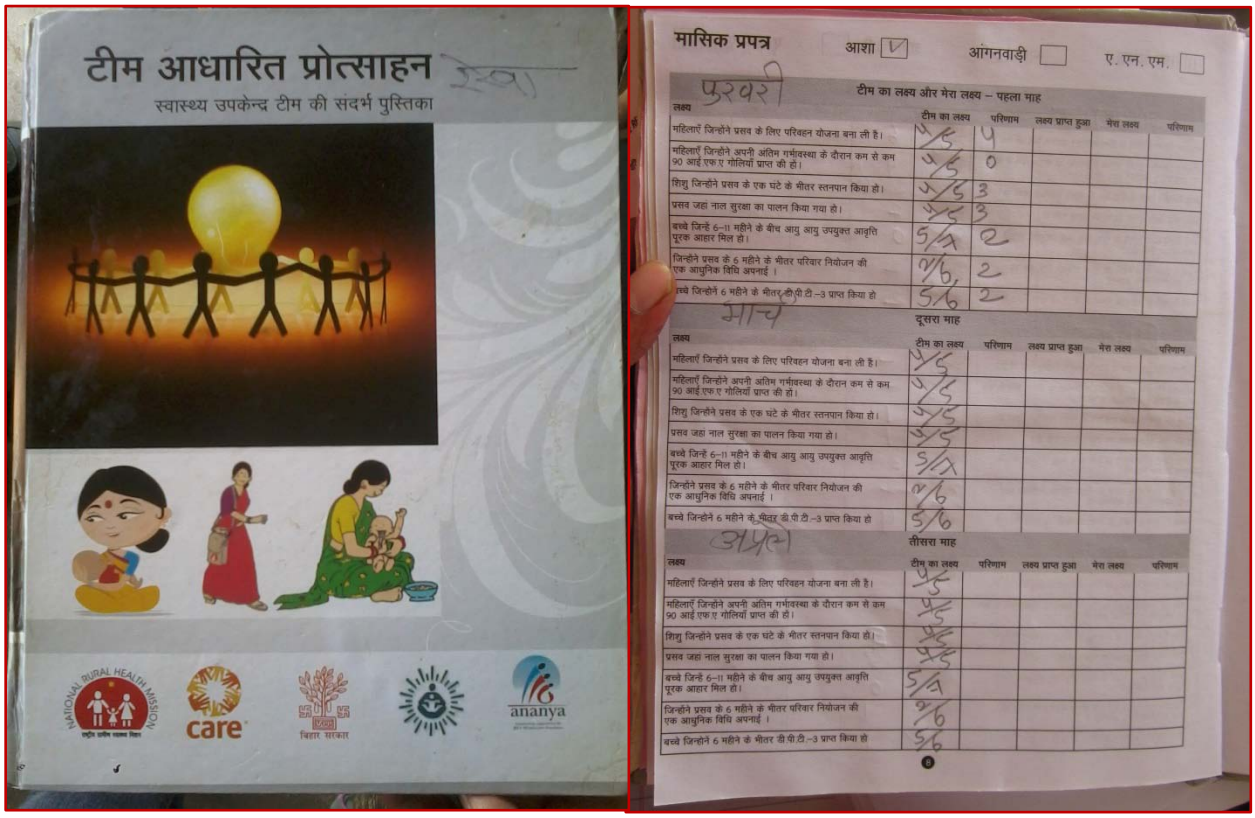
For each of these goals, CARE set specific targets for the percentage of eligible beneficiaries in a catchment area who should have adopted the behavior in a given quarter, using information on the prevalence of the behavior in Begusarai and the extent to which adoption of the behavior might

[1] A subcenter typically covers four to six villages and includes three cadres of FLW: AWWs, ASHAs, and auxiliary nurse midwives (ANMs). There are typically several AWWs and ASHAs per subcenter (with one AWW and one ASHA per village or part of a village), and one or two ANMs in a supervisory position.

reasonably be expected. For all goals except adoption of family planning, the overall targets were set at a fairly ambitious coverage rate of 70 to 80 percent of the beneficiary population. For the adoption of the family planning goal, the target was set lower, at 30 percent, because of the low prevalence of adoption of family planning and the difficulty in promoting this behavior in rural Bihar. These percentage goals were the same across all subcenters in the study, and applied to the combined catchment area of all ASHAs and AWWs in the subcenter (under the supervision of the auxiliary nurse midwife [ANM]). However, each ASHA–AWW pair (typically responsible for a village or a segment of a village served by an Anganwadi Center (AWC) in the subcenter catchment area) was also given the same targets and encouraged to ensure that the targets in their catchment areas were met or exceeded so that the collective subcenter targets could be achieved.

Each ASHA and AWW received a TBGI diary that listed the seven goals and enabled them to record by month how many relevant beneficiaries (pregnant women or women with newborns/infants) were present in their catchment area and how many had attained each particular goal (see Figure I.1). At the end of each quarter, information was tallied on the number of beneficiaries who adopted the behavior in the subcenter's catchment area based on FLW reports, and computed as a percentage of all recorded beneficiaries. The program rewarded *all* the FLWs in the subcenter with small nonmonetary incentives so long as their subcenter met the targets set for at least five of seven goals for that quarter. Thus, the focus of the intervention was not as much setting high targets and penalizing those who could not meet all their targets, but rather to instill a sense of teamwork by encouraging FLWs to collaborate. The quarterly incentives typically consisted of stoves, casseroles, storage containers, or other similar household items. An additional so-called bumper prize (a pressure cooker) and a certificate was given at the end of the year to FLWs in those subcenters that successfully met their targets in all four quarters.

**Figure I.1. Job Aids Used During Home Visits**



The role of the ANM was also an important part of this intervention. The ANM was assigned the task of supervising and tracking progress towards goals achieved by each of the ASHAs and AWWs in her team, as well as assisting them with challenging cases. Further, block-level officials such as the block health manager (BHM), block community mobilizer (BCM), and the child development project officer were also encouraged to visit a couple of villages in subcenters that had achieved their quarterly targets, and the officers were provided with simple guidelines to randomly visit households in these subcenters to cross-verify the results.

At the outset of the program, the CARE team introduced the team-based intervention and the goals to the FLWs in treatment subcenters as part of the monthly subcenter platform meeting. (These subcenter meetings are regular training meetings being implemented as part of the core Ananya program in Bihar). In the initial meeting in each subcenter, the CARE team provided information on the seven goals, how the subcenter targets and targets for each pair of FLWs were set (to ensure that FLWs were comfortable with these targets), the concept of teamwork and working jointly, and the nonmonetary incentives they could receive if they were successful. Besides providing nonmonetary incentives, the program has additional elements to motivate the FLWs. For instance, each subcenter meeting starts with all the FLWs together reciting a pledge that reiterates their mission to assist and guide pregnant women and children in their areas and improve standards of health and nutrition for them. In addition, FLWs in subcenters that meet five of seven goals in every quarter receive a certificate of recognition at the end of the year.

The approach that FLWs used in trying to ensure that households adopted specific behaviors was the same as in the Ananya program: to conduct home visits with beneficiary households and provide them with appropriate messages and services. The FLWs had access to the trainings and tools provided as part of Ananya (in both TBGI treatment and control subcenters). The trainings on health topics were provided as part of the monthly subcenter platform meetings planned and facilitated by CARE staff (and aided by the subcenter ANM) in which maternal and child health topics are discussed, together with the importance of home visits, effective communication, and messaging using the tools and planners provided. In addition, in the treatment subcenters, the FLWs brought their TBGI diaries to the monthly meetings and discussed the progress made in the previous month (or current month) toward attaining the goal for the quarter. At these meetings, the FLWs also could get the support of the ANMs or other FLWs in terms of how to address particular challenging issues.

The initial operational plan of the program was finalized after CARE conducted a formative study in spring 2012. As part of this study, CARE held focus groups and consulted with FLWs to help assess whether an outcome-based (rather than a process-based) incentive program was likely to be feasible and whether this was a strategy the FLWs could adopt, and to identify the types of nonmonetary rewards that might appeal to the FLWs. The actual intervention was formally launched in August 2012 in five blocks (38 subcenters) in Begusarai.

## II. STUDY DESIGN, DATA, AND ANALYTIC APPROACH

We used a rigorous RCT design to study the effectiveness of the TBGI intervention, and conducted a process study to learn about its implementation. In this chapter, we briefly describe our study design, data collection, and analysis approach for the RCT, as well as our approach to the qualitative data collection.

## A. Design for the RCT

To provide rigorous, highly credible evidence on the impact of the TBGI intervention, we conducted a clustered RCT. Because the intervention was implemented at the subcenter level, we randomly assigned subcenters in five blocks in Begusarai district to a treatment group, which received the TBGI intervention, or to a control group, which did not receive the TBGI interventions but had access to other services available in the communities (including all other elements of the Ananya program). Randomization ensures similar groups of FLWs and beneficiaries in the treatment and control communities prior to the start of the intervention, on average, except that only those in the treatment subcenters are exposed to the intervention. As a result, we are able to attribute any observed differences in outcomes over time for beneficiaries or FLWs to the causal effect of the TBGI intervention.

Because FLWs and households in the treatment and control groups received non-TBGI Ananya program interventions that were being implemented simultaneously across Begusarai, the RCT was designed to measure the *value-added* of TBGI beyond these other Ananya interventions. It was *not* designed to measure the impact of TBGI introduced in isolation. Indeed, many of the outcomes that we examined as part of this study—especially health behaviors and practices among beneficiaries—showed strong positive trends between the baseline and endline, even in the control group.[2] The TBGI impacts should, therefore, be interpreted in a context in which large background changes occur simultaneously, rather than in a static context.

CARE selected Begusarai as the district in which the TBGI intervention would take place based on the willingness of district-level government officials to support the study. To limit implementation costs, the intervention and RCT were focused in five blocks (of 18 total blocks in Begusarai). We selected the five blocks to represent a range of sizes and geographies, and we excluded blocks that CARE's ground team had qualitatively determined to be atypical (such as those in which government health officials were not in place). Our statistical power calculations suggested that we required about 80 subcenters for the study, and the five selected blocks were adequate to provide the sufficient number of subcenters. (Our power calculations suggested that with a sample of 80 subcenters and 20 women per subcenter, we would be able to detect impacts of 6 to 10 percentage points in the behaviors targeted by the intervention.)

We randomized all the subcenters in the selected blocks into equal-sized treatment and control groups using a stratified random assignment procedure based on the number of Anganwadi Centers (AWCs) served by the subcenter (a proxy for the size of the population served). The stratification helped to ensure that the treatment and control groups were balanced by the size of the population served, and to reduce variance (and hence improve statistical power) in the analysis (see Appendix A for more details).

[2] Although we cannot precisely disentangle these trends in control group beneficiaries from baseline to endline from other general trends unrelated to Ananya that might lead to change in health outcomes, their large magnitude suggests that the non-TBGI Ananya program interventions were likely associated with changes in practices for both the treatment and the control group beneficiaries.

## B. Sample and Data for the RCT

In both the TBGI treatment and control subcenters, we conducted interviews with FLWs and beneficiaries who had given birth in the past year. The TBGI intervention—and the Ananya program as a whole—envisages the ASHAs, AWWs, and ANMs working together in an integrated fashion to interact with and improve the health of mothers and young children in the communities. We therefore sought to obtain information from these FLWs on aspects of working as a team, their attendance and participation in subcenter meetings, and home visits they had conducted and services they provided to beneficiaries. We also gathered information from women who had given birth in the past year on their behaviors and practices in the health areas targeted by TBGI—including antenatal care and delivery preparation, postpartum care, child immunizations, complementary feeding, and family planning—as well as their interactions with the FLWs.

We conducted two rounds of data collection from both FLWs and beneficiaries: a baseline (in May and June 2012, before the TBGI intervention was introduced) and an endline a little over a year later.[3] The baseline surveys were intended to enable us to verify the comparability of the treatment and control groups at baseline, and to provide control variables to improve the precision of our impact estimates. The endline, conducted in August and September 2013, was used to determine the impacts of the TBGI intervention after about one year of exposure.

Because there was no readily available sample frame available to identify women who had given birth in the past year, we conducted a household listing to identify the right set of beneficiaries to survey. Conducting a full listing of all households in the subcenter would have been prohibitively expensive; therefore, we randomly selected two villages (or village segments) per subcenter in which to conduct the listing, defining segments in such a way that we identified about 20 eligible beneficiaries per segment (see Appendix A). We surveyed women in the same communities at baseline and endline. However, the household beneficiaries included in the baseline and endline samples were largely different, because a different group of women gave birth in the 12 months before each survey. Nevertheless, baseline data enabled us to check that the randomization was successful in creating comparable groups at baseline, and it provided baseline subcenter-level control variables to improve the precision of our impact estimates at endline.

For the FLW surveys, we wanted to interview about four ASHAs and four AWWs per subcenter (our power calculations suggested that this combined sample would enable us to detect impacts of 7 to 12 percentage points in key FLW outcomes). For our baseline FLW surveys, we used a list of subcenter ASHA and AWWs provided by CARE; we interviewed ASHAs and AWWs who served beneficiaries in the villages (or village segments) selected for the beneficiary surveys, and supplemented this with an additional random sample of ASHAs and AWWs from the subcenter (see Appendix A for details). At endline, we attempted to survey all the ASHAs and AWWs identified at baseline, plus any new ASHAs and AWWs serving households in the sampled communities. Because each subcenter had only one or two ANMs, we attempted to survey *all* the ANMs in the subcenters included in our sample at baseline and endline, respectively.

The response rates to our surveys were generally very high (Table II.1). Focusing on the endline surveys, about 95 percent of households responded to the listing survey and about 89 percent of

[3] Sambodhi Research and Communications conducted both the baseline and the endline data collection.

eligible women responded to the beneficiary survey (this combined response rate is about 84 percent). This yielded a total sample of 1,607 completed interviews for the endline beneficiary-level impact analysis. These numbers exclude four subcenters (two in treatment and two in control) that, because of flooding, could not be surveyed at endline and were therefore dropped from the impact analysis.[4] Among FLWs, the endline response rates were about 80 percent for AWWs, 77 percent for ASHAs, and 86 percent for ANMs—yielding a total FLW sample size of 638. The response rates were similar in the treatment and control groups, suggesting that differential nonresponse between the treatment and control groups is unlikely to bias our estimates.

**Table II.1. Endline Sample Sizes and Response Rates**

| | Control (36 subcenters) | | Treatment (36 subcenters) | | Total Combined (72 subcenters) | |
|---|---|---|---|---|---|---|
| Survey | Response Rate (percentage) | Sample Size | Response Rate (percentage) | Sample Size | Response Rate (percentage) | Sample Size |
| **Households** | | | | | | |
| Listing Survey | 95.2 | 7,860 | 92.4 | 7,719 | 94.4 | 15,514 |
| Beneficiary Survey | 89.4 | 809 | 87.6 | 798 | 88.5 | 1,607 |
| **FLWs** | | | | | | |
| AWW | 77.3 | 136 | 81.3 | 135 | 80.4 | 275 |
| ASHA | 74.9 | 137 | 81.7 | 138 | 77.0 | 271 |
| ANM | 92.7 | 51 | 78.8 | 41 | 86.0 | 92 |

Source: Ananya TBGI RCT Endline

Note: ANM sample sizes include four additional subcenters (two in treatment and two in control) in which households, ASHAs, and AWWs could not be surveyed due to flooding.

## C. Analytic Approach to Estimating Impacts

Because randomization should ensure that the treatment and control groups are similar in all respects other than receipt of the intervention, impacts can be estimated simply by computing the difference in mean outcomes between the two groups at endline. However, we instead estimated impacts in a regression framework, which enabled us to explicitly account for the method of randomization by including stratum-level indicators (see Appendix A for details). By controlling for additional individual and baseline subcenter-level characteristics, we were also able to reduce the variance in the outcome (and hence increase statistical power) and control for differences between the treatment and control groups that could have arisen by chance.[5]

In our regressions, we included both individual- and subcenter-level covariates, respectively, that could be related to the outcome of interest. For the analysis of beneficiary outcomes, our individual-level covariates included household demographics, such as whether a woman belongs to a scheduled caste or tribe, is Muslim, the number of children, age, literacy, and SES quartile; the subcenter-level covariates included the subcenter-level means of these same outcomes, when

[4] These subcenters were also dropped from the ASHA and AWW impact analyses because we were unable to survey ASHAs and AWWs there. However, we were able to survey ANMs in these subcenters, and included them in the analysis.

[5] As a robustness check, we compared our results with those from simple unadjusted treatment–control comparisons; the results were largely similar.

available.[6] We weighted all beneficiary-level regressions to account for differing sampling probabilities and to ensure treatment–control balance within random assignment strata—the estimated impacts can therefore be interpreted as the impacts for the average beneficiary in the treatment subcenters. We adjusted all standard errors to account for the correlation in outcomes among beneficiaries linked to the same subcenter.

We used a similar regression framework to determine the impact of the interventions on outcomes for ASHAs and AWWs and include both individual FLW and subcenter characteristics. In all our FLW analyses, we pooled the data collected from ASHAs and AWWs.[7] Subcenter-level covariates included the share of beneficiaries served by the subcenter who are literate, Muslim, in a scheduled caste or scheduled tribe, and first-time mothers. FLW-level covariates characteristics include controls for the worker's age, scheduled caste or scheduled tribe status, religion, education level, and an indicator for whether the FLW serves a village less than one hour's travel time from her subcenter. We estimated a more parsimonious model for our ANM sample, which did not include an extensive set of covariates. Given the small sample size of this group, we do not have the statistical power to distinguish impacts; hence, the ANM results should be thought of as providing descriptive context for the ASHA, AWW, and beneficiary results. As for the beneficiary analysis, we weighted all regressions so that they represent the average FLW (either the average ASHA/AWW or the average ANM, depending on the analysis), and adjusted for clustering of standard errors at the subcenter level.

## D. Qualitative Process Study

To learn about program implementation and to inform the RCT findings, we conducted a process study in which we gathered qualitative information on the implementation of TBGI. We collected these data as part of field visits held at the end of April and early May 2013, at which time the TBGI intervention had just completed its third quarter (February to April 2013).

As part of the process visits, we conducted interviews with CARE program staff at headquarters and at the district level to learn about their vision for the TBGI program and how it was implemented. In particular, we asked about inception of the TBGI intervention, planning and rollout for the intervention, perceptions of staff regarding implementation successes and challenges, and their views about the scalability of the intervention.[8]

We also conducted interviews with the FLWs and beneficiaries to understand program implementation from their perspectives. In particular, we visited three of the five blocks in which the program was implemented and conducted semistructured interviews with 24 FLWs (8 ANMs, 9 ASHAs, and 7 AWWs) working at 10 subcenters. We reviewed their TBGI planners and went to a subcenter platform meeting. We asked FLWs about the monthly meetings held at the subcenter level; how they coordinate their home visits; their use and knowledge about job aid tools and registers; and information related to the TBGI intervention, such as how they were informed about

[6] Because we made some changes to the beneficiary survey between baseline and endline to better capture aspects of the intervention, some of the outcomes analyzed at endline were not measured in the baseline survey.

[7] We also analyzed several key outcomes for ASHAs and AWWs separately, but found few results that varied across cadres.

[8] We also met with the BHM in one of the block primary health centers (PHCs) to learn about their awareness and understanding of the program and its scalability.

the program; the training process; how the goal and targets for the program were set; and how the FLWs perceived the incentives that were provided to them for the accomplishment of these goals. Our team also conducted interviews with 18 household beneficiaries in the TBGI catchment area, including pregnant women, mothers with newborns, and mothers with children from 6 to 11 months of age to obtain their perspectives on home visits and what topics were discussed. We systematically triangulated and synthesized the findings from the process study to answer key questions related to program implementation and summarize these in subsequent chapters.

# III. IMPLEMENTING THE TBGI INTERVENTION

In this chapter, we summarize our findings on how CARE staff implemented the TBGI intervention and how FLWs in the treatment subcenters understood and implemented it. These findings are drawn from two sources: (1) semistructured interviews we conducted with CARE staff and FLWs in treatment subcenters, which gathered information on their experiences with the program and key successes and challenges in implementation; and (2) a module of the ASHA, AWW, and ANM surveys that asked questions on the specifics of the intervention for those in the treatment subcenters, including the guidance and instruction received and their perceptions and experiences related to key activities. While the semistructured qualitative interviews we conducted were not designed to yield quantitative estimates of perceptions or experiences (given the small sample sizes involved), they did enable us to draw some general conclusions based on commonalities and differences in responses across the respondents in our sample. Below, we triangulate the qualitative data with the quantitative survey data we collected on treatment subcenter FLWs for our randomized controlled trial.

## A. FLWs' Understanding of the TBGI Intervention and its Key Elements

Overall, ASHAs and AWWs understood the essence of the program and reported ease in achieving the targets set for them. CARE's collaborative effort of developing and implementing the program with input from the FLWs was evident, as a majority of the FLWs we interviewed were able to describe the key elements of the program and explain how it worked.

Although most FLWs knew that CARE set targets for select outcomes and were aware of how many they had to meet to receive the incentives, few were able to explain precisely how the numerical targets were calculated. The process of setting numerical targets involved identifying the number of target beneficiaries for each outcome (which the FLWs were usually able to report), and then using the percentage of that group that had to attain the outcome (for example, 70 or 80 percent) to calculate the numerical target. Despite CARE's explanations, and provision of simple multiplication tables, this was a difficult calculation for many FLWs. For example, as seen in Table III.1, although 59 percent of FLWs were able to report the number of households that would have to arrange for transportation to the health facility for delivery in order for them to meet this target and only 3 percent of FLWs reported a target that was not correct or within one beneficiary of the correct number, a significant minority reported that they did not calculate this goal themselves. According to our surveys with the FLWs, about 65 percent reported that they had determined the transportation goal themselves, 20 percent did so with assistance, and 15 percent said they were given the target. ASHAs and AWWs also received support with tracking progress toward these targets. For instance, 73 percent reported that someone from CARE reviewed their diary in the past month.

Our qualitative interviews suggest a similar pattern. During our field visits, some ASHAs and AWWs were able to show and explain to us the multiplication tables that CARE had given them to use while conducting the calculations. However, the more common response was that they relied on CARE's staff at the subcenter meetings to calculate their targets each quarter. ANMs themselves were sometimes not able to articulate clearly how the target percentages were calculated and relied on the CARE team to help with the process of calculating and verifying whether the targets were met. In addition, according to both ANM reports and our own observations, some ASHAs and AWW's (particularly the less literate) faced difficulties in filling out their TBGI diaries to track their progress towards their targets.[9] Helping FLWs understand how to set specific targets in a simple manner and to record progress on their own without the support of CARE staff will be an important consideration if the program scales up.

**Table III.1. Understanding Targets and Tracking Progress Toward Them**

| | Mean |
|---|---|
| **Calculating Targets (example: transportation goal)** | |
| Knew the Correct Target (number of beneficiaries) | 58.7 |
| Knew the Correct Target (plus or minus one beneficiary) | 96.8 |
| Determined Target | |
| By self | 65.1 |
| With someone's help | 19.8 |
| Someone gave target | 15.1 |
| **Understanding When Goals Are Met (example: 90+ IFA tablet goal)** | |
| After giving tablets | 33.2 |
| After asking if received tablets | 25.9 |
| After asking if consumed tablets | 71.9 |
| After asking to see empty strips | 52.4 |
| After asking about stool color | 41.2 |
| Other FLW is responsible for this target | 1.6 |
| After seeing the woman's tongue's color | 4.5 |
| **Tracking Progress Toward Targets** | |
| Observed TBGI Diary | 92.3 |
| Has Own TBGI Diary | 84.2 |
| Has Shared TBGI Diary | 11.0 |
| Diary Reviewed in Past Month by | |
| Anyone | 94.5 |
| ANM | 41.7 |
| Someone from CARE | 72.6 |

Source: Ananya TBGI RCT endline survey, treated subcenters only.

Note: The joint sample size of ASHAs and AWWs ranges from 261 to 273. All means are weighted.

[9] As part of our process study visits, we observed that FLWs in treatment areas seemed to keep their home visit planners more up to date than FLWs in nonexperimental communities (a home visit planner is another Ananya tool provided by CARE to facilitate home visit scheduling and tracking). This could be because CARE staff asked to see these planners in order to help calculate the target number of beneficiaries. Our surveys of FLWs, in which we found that our investigators observed a larger fraction of treatment subcenter FLWs filling out home visit planners relative to FLWs in control subcenters, corroborated these findings.

In both our semi-structured interviews and in the surveys, FLWs articulated a varied understanding of what accomplishing each of the seven goals meant. For example, with regard to the goal related to IFA tablets, although some FLWs interpreted the distribution of 90 IFA tablets to pregnant women in their catchment area as having met the goal, others adopted approaches to verify if the beneficiaries had consumed the tablets. In our FLW surveys, 35 percent felt they achieved the IFA goal if they gave a woman the tablets, but 70 percent thought it was necessary to ask the woman if she consumed the tablets. A little under half (43 percent) of the FLWs also felt it was necessary to ask a woman about the color of her stools (which turn black upon consumption of IFA tablets). There was a similar range of variation across the other goals (see Appendix Table B.1).

## B. Motivation and Teamwork

The ASHAs and AWWs we interviewed were excited and enthusiastic about the program and felt that it motivated them to do more. For instance, 96 percent of all respondents reported that reciting the pledge that CARE had crafted—in which they commit to improving health care for women in children in their area—motivated them a lot (see Figure III.1 and Table III.2). This motivation was also evident in the sense of joint ownership of the program among FLWs; 80 percent of ASHAs and AWWs reported both they and their cadre partner shared responsibilities for all households in their catchment area; 75 percent reported sharing responsibility for all goals (see Table III.2).

**Figure III.1. FLWs Signing Pledge**

Source: CARE, "Background," Team Based Goals and Incentives (blog), February 2013. Available at [http://tbgibeg.blogspot.com/2013/02/background.html]. Accessed February 13, 2014.

**Table III.2. Motivation and Teamwork**

| | Mean |
|---|---|
| Reciting Pledge Motivates You a Lot | 96.0 |
| Share Responsibility for All Households with ASHA/AWW | 79.7 |
| Share Responsibility for All Goals with ASHA/AWW | 74.9 |
| Meet with ASHA to Update Each Other on Progress Toward Goals | |
| Every day | 9.7 |
| A few times a week | 60.1 |
| Once per week | 19.9 |
| A few times per month | 4.8 |
| Never or only at subcenter meetings | 5.5 |
| ANM Provides Advice on Meeting Targets | |
| Most of the time | 59.5 |
| Sometimes | 35.7 |
| Very few times | 4.7 |

Source: Ananya TBGI RCT endline survey, treated subcenters only.

Note: The joint sample size of ASHAs and AWWs is 273. All means are weighted.

Accordingly, ASHAs and AWWs tended to work as a team to cover the households, either by conducting joint visits or by dividing the beneficiaries between themselves. They often discussed goals and progress; 70 percent of ASHAs and AWWs reported meeting a few times per week or more to talk about their progress toward targets and 20 percent more met on a weekly basis (see Table III.2). The monthly subcenter meetings reinforced this trend. They strongly encouraged the ASHAs and AWWs to work closely with each other to meet the goals.

Our process study had similar findings. Most ASHAs and AWWs we interviewed reiterated that they relied on each other to complete home visits. They explained that their FLW counterparts assisted them by conducting home visits on days they could not. We did encounter one or two ASHAs or AWWs who felt that there had been no change in interactions with their counterparts or that their workloads had increased because the other FLW did not do her part. However, only a small minority of the respondents expressed these views.

In addition to ASHAs and AWWs, the ANM is also a key member of the subcenter team. We explored through both our process study interviews and FLW surveys the role of the ANM in facilitating the work conducted by ASHAs and AWWs. CARE's vision was that the ANM would provide feedback on the progress ASHAs and AWWs have made in accomplishing their goals and to review their registers for accuracy. In addition, she is expected to guide them on what information to share with the households and how to communicate key messages. The ASHAs and AWWs we surveyed indicated that they reached out to their ANMs for assistance in cases in which beneficiaries were reluctant to follow a practice. More than half (60 percent) said that the ANM provided advice on how to meet targets "most of the time," whereas 36 percent reported that the ANM sometimes provided this input (see Table III.4). ASHAs and AWWs we interviewed for the process study provided similar information. They reported that their ANM typically provided advice on how to communicate effectively with target households, and sometimes even visited the households herself to convince them of the importance of a practice.

## C. FLWs' Perceptions of the TBGI Goals and Targets

FLWs reported in general that it was not difficult for them to achieve the targets, especially because they had a three-month window to do so. They said they achieved these targets through

home visits, conversations while accompanying women for delivery, or discussions during immunization days. They typically indicated that these goals were within the general scope of their work or what they were expected to do as part of their job. They did not regard the accomplishment of these goals as new or extra work. Several also mentioned how their work was almost easier now given that their visits with households had a more precise focus. They felt that the program's emphasis on specific goals helped them target key beneficiaries and facilitated their delivery of appropriate messages to households during home visits.

FLWs also reported that some goals were easier to achieve than others. In our survey, we asked FLWs to identify the easiest and hardest goals for them to achieve. As seen in Table III.3, there was substantial dispersion in the goal identified as easiest to achieve, though FLWs most commonly reported ease in achieving outcomes related to birth preparedness and antenatal care. About a third (32 percent) felt the easiest goal to achieve was immediate breastfeeding, 23 percent cited transportation, and 19 percent noted DPT-3. There was great consistency in the goal that was mos difficult to achieve: most FLWs (88 percent) felt it was hardest to achieve the family planning goal. Sub-center meetings offered guidance to FLWs on how to meet goals and persuade households to adopt the program's target practices. More than two-thirds (70 percent) of these FLWs reported receiving advice on achieving their goals at most subcenter meetings and 29 percent reported getting this input sometimes.

**Table III.3. Achieving Goals**

| | Mean |
|---|---|
| Easiest Goal to Achieve | |
| Transportation | 22.5 |
| IFA tablets | 12.1 |
| Immediate breastfeeding | 31.6 |
| Cord care procedures | 5.8 |
| Complementary feeding | 6.7 |
| Family planning | 2.0 |
| DPT-3 | 19.3 |
| Hardest Goal to Achieve | |
| Transportation | 3.1 |
| IFA tablets | 2.0 |
| Immediate breastfeeding | 1.9 |
| Cord care procedures | 0.9 |
| Complementary feeding | 2.7 |
| Family planning | 87.7 |
| DPT-3 | 1.8 |
| Receive Advice on Meeting Goals at Subcenter Meetings | |
| Most of the time | 69.5 |
| Sometimes | 29.3 |
| Very few times | 1.3 |

Source: Ananya TBGI RCT endline survey, treated subcenters only.

Note: The joint sample size of ASHAs and AWWs is 273. All means are weighted.

Our process study interviews with a small number of pregnant women and new mothers confirmed the above findings on FLWs' reactions to the TBGI goals. Beneficiaries reported more home visits by the FLWs than in the past and reported implementing important delivery preparation, antenatal care, and newborn care practices. For example, many beneficiaries reported making arrangements for transportation and receiving IFA tablets. Most respondents also said they initiated immediate breastfeeding and cord care after the birth of their children. Adherence to complementary feeding and family planning practices was lower. Few mothers with children

6 months of age had initiated complementary feeding at the time of the qualitative interviews, and adoption of family-planning methods within six months of delivery was rare.[10]

## D. FLWs' Perceptions of the Nonmonetary Incentives

The gifts provided as incentives for the FLWs were identified and selected based on recommendations they made during the formative research conducted by CARE before rolling out the intervention. These gifts included items such as a nonstick *tawa* (griddle), hot-pot (a container to keep food warm), blanket, gas *chulha* (stove), electric iron, flashlight, or other small household items (see Figure III.2). The FLWs also received an annual bumper prize (pressure cooker) and a certificate if their subcenter met five of the seven targets in all four quarters.

**Figure III.2. FLWs Receiving Incentives for the First Quarter**

Source: CARE, "Background," Team Based Goals and Incentives (blog), February 2013. Available at [http://tbgibeg.blogspot.com/2013/02/background.html]. Accessed February 13, 2014.

In both our surveys and interviews, most FLWs expressed satisfaction with the quality of gifts they received, though a few mentioned that the quality was low. Table III.4 details these responses from the former. Almost three-quarters (73 percent) of ASHAs and AWWs who received the

[10] See Chapter V for more detail on beneficiary-level outcomes from our quantitative beneficiary-level survey.

nonmonetary incentives were very satisfied with their gifts and only 8 percent were somewhat or very dissatisfied. Most also reported using their gifts, with 61 percent saying they did so on a daily basis. Despite this, FLWs reported a preference for receiving certificates over gifts. The vast majority (89 percent) said they preferred the former over the latter, though this might reflect the respondents' tendency to give socially desirable responses.

**Table III.4. Reactions to Nonmonetary Incentives**

| | Mean |
|---|---|
| Satisfaction with Gifts (given or received) | |
| Very satisfied with gifts | 73.1 |
| Somewhat satisfied with gifts | 18.9 |
| Somewhat dissatisfied with gifts | 6.1 |
| Very dissatisfied with gifts | 1.8 |
| Use Most Recent Gift (if any) | |
| Every day | 61.0 |
| Sometimes | 19.1 |
| Never | 19.9 |
| Prefer Gift Over Certificate | 11.0 |
| Prefer Certificate Over Gift | 89.0 |

Source: Ananya TBGI RCT endline survey, treated subcenters only.

Note: The joint sample size of ASHAs and AWWs ranges from 268 to 273. All means are weighted.

FLWs also mentioned other less tangible incentives that the TBGI intervention provided. They noted, for example, that the efficient roll-out of the incentives each quarter motivated them to do better, as did the frequent acknowledgement of their hard work. A few also mentioned how the program had instilled a sense of competition among FLWs from different subcenters, which had positively affected their work. One FLW mentioned, however, that she was apprehensive about her ability to achieve the targets and concerned that her reputation might be affected given that every FLW's progress was known publicly.

## E. Monitoring of TBGI Activities

CARE envisions that government officials will monitor FLW activities under the TBGI intervention. Specifically, block-level primary health center (PHC) staff such as BHMs or BCMs would conduct a small number of random home visits on a quarterly basis to validate the results reported by the FLWs. These officials are expected to randomly select a small number of AWCs each quarter from a subcenter that has met its targets. They would visit six households in that area to verify the adoption of practices that the subcenter had reported that its households had started to follow. We did not hear about block-level staff conducting verification visits in our semistructured interviews. FLWs did mention that other people such as CARE officials or other external parties sometimes accompanied them on home visits, but these visits seemed related mostly to observations rather than back-checks to verify reports.

It is possible that, given our relatively small qualitative sample size and the fact that verification visits were conducted in only a few AWCs, we simply did not interview FLWs whose work had been monitored by block-level officials. However, our survey data set also suggests that CARE took the lead on verifying reports in the TBGI diaries. 12 percent of ASHAs and AWWs reported having these numbers verified by a BHM or BCM, 36 percent by an ANM, and 77 percent by CARE staff.

Despite these findings, it is possible that the nature and scope of verification by block-level officials might change over time if the Government of Bihar scales up the intervention. For example, verification might be formally integrated into the responsibilities of block-level officials, which was not the case during the study period.

# IV. RESULTS FROM THE RCT: FLW SURVEYS

The logic of the TBGI intervention suggests that for it to have affected beneficiary-level outcomes, it must first have resulted in changes in FLWs' behavior and interactions with mothers and children. Therefore, we first discuss the results from our analysis of FLW outcomes. Although these FLW-reported outcomes provide valuable evidence on the impacts of TBGI, it is important to keep in mind that FLWs might have some incentive to over-report on the services they provided. When possible, we also reviewed written records that FLWs kept and describe beneficiaries' own reports of FLWs' behaviors to triangulate our results (see Chapter V). In addition, we asked the ANMs about the ASHA and AWW activities, and asked each ASHA and AWW to report on the behavior of other FLWs in her community (as well as about whom she considers to be part of her team). Although the impact evaluation results do not always align perfectly across the ASHAs, AWWs, ANMs, and beneficiary surveys, they point to the same broad conclusions.

We begin this chapter by describing the baseline characteristics of FLWs in the treatment and control subcenters. Although random assignment should ensure that our results provide valid measures of the impact of TBGI, an examination of the equivalence of baseline characteristics provides further assurance that there were no large differences in outcomes for the treatment and control groups before the intervention, or large differences in the characteristics of these groups. We then consider the impact of TBGI on how FLWs work together to serve mothers and children and the services they provide. All tables for this chapter and the next appear at the end of the report.

## A. Baseline Equivalence of Treatment and Control FLWs

To assess whether the FLWs in treatment and control subcenters were similar before the implementation of TBGI, we examined baseline measures of outcomes (Table IV.1). Although changes in the FLW surveys between baseline and endline made it difficult to directly compare outcomes at these points, the outcomes at baseline still capture measures similar to those we focus on at endline.[11]

Overall, our analyses indicate that our samples of FLWs were similar before the TBGI intervention based on both their demographic characteristics and their activities and services provided at baseline. Before the intervention, ASHAs and AWWs at treatment and control subcenters participated similarly in subcenter meetings (with more than 70 percent attending three or more in the past three months), and they reported providing similar coverage of households (Table IV.1). Further, topics they reported discussing during home visits were similar for treatment

[11] We changed the FLW questionnaire considerably from baseline to endline to better capture aspects of the intervention as it was rolled out in practice, and to reflect lessons learned from our analysis of the baseline data.

**Table IV.1. Baseline Outcomes for Treatment and Control FLWs (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **ASHA/AWW Sample** | | | | |
| Participated in 3 or More Subcenter Meetings in Past 3 Months | 72.2 | 76.1 | 3.9 | 0.448 |
| **Home Visits** | | | | |
| Have Observed Home Visit Planner | 27.0 | 28.9 | 1.9 | 0.752 |
| Visited All Pregnant Women in Catchment Area in Their Third Trimester | 58.2 | 57.5 | -0.7 | 0.892 |
| Visited All Women Who Delivered in Past Three Months Within One Week of Birth | 49.5 | 46.4 | -3.1 | 0.570 |
| Visited All Mothers of Children in Catchment Area Around 6 Months | 32.3 | 26.1 | -6.2 | 0.221 |
| Used Materials or Job Aids During Home Visits Last Month | 52.0 | 56.0 | 4.0 | 0.481 |
| **Working in a Team** | | | | |
| Times Met with ASHA/AWW in Past 7 Days | 2.14 | 2.15 | 0.10 | 0.931 |
| Coordinate Home Visits with ASHA/AWW | 68.2 | 73.9 | 5.7 | 0.288 |
| Percentage Agreeing or Strongly Agreeing:[a] | | | | |
| Goals are clear | 76.0 | 83.2 | 7.3* | 0.088 |
| Team works together to set goals | 75.0 | 83.3 | 8.2* | 0.051 |
| Team cares about meeting goals | 77.9 | 85.1 | 7.2* | 0.091 |
| Team can reach most targets | 79.3 | 87.6 | 8.3** | 0.046 |
| **Sample Size** | **316** | **303** | | |
| **ANM Sample** | | | | |
| **Subcenter Meetings** | | | | |
| Participated in 3 or More Meetings in Past 3 Months | 69.1 | 70.5 | 1.4 | 0.850 |
| Ran Last Subcenter Meeting Alone | 3.6 | 9.7 | 6.1 | 0.266 |
| Ran Last Subcenter Meeting with Someone Else | 51.8 | 46.8 | -5.0 | 0.624 |
| **Services** | | | | |
| Performed Antenatal Care Checkups in the Past Calendar Month | 96.4 | 96.0 | -0.4 | 0.922 |
| Identified Newborn with Danger Signs | 7.1 | 12.0 | 4.9 | 0.386 |
| Assisted Woman with Serious Postpartum Complications to See Doctor in Past Calendar Month | 12.5 | 14.1 | 1.6 | 0.784 |
| Distributed Contraception in Past Calendar Month | 80.4 | 76.8 | -3.5 | 0.653 |
| **Supervision** | | | | |
| Times Per Month Meet with Each ASHA | 4.46 | 4.30 | -0.15 | 0.713 |
| Times Per Month Meet with Each AWW | 1.14 | 1.25 | 0.11 | 0.788 |
| **Skills** | | | | |
| Percentage Needing Additional Skills for Job | | | | |
| Any | 85.7 | 90.1 | 4.4 | 0.431 |
| Record-keeping | 41.1 | 36.8 | -4.2 | 0.635 |
| Maternal health issues | 50.0 | 49.4 | -0.6 | 0.949 |
| Newborn health issues | 53.6 | 64.3 | 10.8 | 0.241 |
| Communicating with families | 23.2 | 32.4 | 9.2 | 0.363 |
| **Sample Size** | **56** | **48** | | |

Source: Ananya TBGI RCT Baseline and Endline.

Notes: Item-specific nonresponse might limit the sample size for some comparisons. We adjusted the treatment mean for differences by stratum and FLW cadre using a regression including stratum by cadre-fixed effects.

[a] Response options are strongly disagree, somewhat disagree, somewhat agree, and strongly agree.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

and control groups, as were their self-reports on the degree to which they met and coordinated with one another. For example, ASHAs and AWWs in treatment areas met an average of 2.15 times per week and those in control areas met an average of 2.14 times per week. As an additional check of comparability between the treatment and control groups, we compared the background characteristics of the FLWs we surveyed at endline, which were unlikely to have changed as a result of the intervention. These characteristics were largely similar, both for the combined ASHA/AWW sample and the ANM sample (Table IV.2).

**Table IV.2. Background Characteristics of Treatment and Control FLWs at Endline (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **ASHA/AWW Sample** | | | | |
| Religion | | | | |
| Hindu | 95.4 | 94.3 | -1.0 | 0.685 |
| Caste (if Hindu) | | | | |
| SC/ST | 13.0 | 11.2 | -1.8 | 0.553 |
| OBC | 59.9 | 55.8 | -4.1 | 0.448 |
| Education Below 8th Standard (ASHAs only) | 2.3 | 0.6 | -1.7 | 0.448 |
| Education Below 10th Standard (AWWs only) | 6.0 | 1.6 | -4.4 | 0.110 |
| **Sample Size** | **273** | **273** | | |
| **ANM Sample** | | | | |
| Religion | | | | |
| Hindu | 92.2 | 0.973 | 5.1 | 0.230 |
| Caste (if Hindu) | | | | |
| SC/ST | 8.5 | 4.6 | -3.9 | 0.450 |
| OBC | 17.0 | 29.6 | 12.6 | 0.182 |
| Attended College, Took College-Level Courses, or Received Diploma | 43.1 | 59.1 | 15.9 | 0.129 |
| **Sample Size** | **51** | **41** | | |

Source: Ananya TBGI RCT Baseline and Endline.

Note: Hindu subsample contains 260 control ASHAs/AWWs, 257 treatment ASHAs/AWWs, 47 control ANMs, and 40 treatment ANMs. The ASHA subsample contains 137 control and 138 treatment ASHAs. The AWW subsample contains 136 control and 135 treatment AWWs. The treatment mean was adjusted for differences by stratum and FLW cadre using a regression including stratum by cadre-fixed effects.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

Thus, overall, we conclude that random assignment led to the creation of equivalent treatment and control groups. However, there were statistically significant differences in one set of outcomes at baseline: those related to ASHAs' and AWWs' perceptions of setting and meeting goals (Table IV.1). In particular, ASHAs and AWWs in treatment subcenters were significantly more likely to report that their goals were clear and that their teams worked together to set goals, cared about meeting goals, and could reach their targets even at baseline(although only the last of these was significant at the 5 percent level). Given that the TBGI intervention did not begin until after the

baseline survey (and that all the pre-baseline formative work took place in nonintervention subcenters), these differences are likely to be unrelated to the intervention and simply reflect the fact that when we examine a large number of indicators, we would normally expect to see at least one or two significant differences by chance.[12] We examined the sensitivity of our findings to this existing difference by adding measures of perceptions of teamwork from baseline to our regression models. This did not affect our endline impact estimates in any meaningful way, which indicates that our impact estimates reported are robust.

## B. Impacts of TBGI on Key Outcomes: Teamwork and Coordination

The most proximal aim of the TBGI intervention was to increase motivation, teamwork, interactions, and joint problem-solving among FLWs, to lead in turn to more and better services to households and eventually to improvements in maternal and child health outcomes. Table IV.3 examines some key indicators related to attendance at subcenter meetings and topics discussed there, as well as indicators of teamwork and coordination as reported by ASHAs/AWWs and ANMs.

**Table IV.3. Teamwork and Coordination of FLWs (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **ASHA/AWW Sample** | | | | |
| **Subcenter Meetings** | | | | |
| Attended Three or More Subcenter Meetings in Past Three Months | 82.7 | 93.9 | 11.2** | 0.000 |
| ASHAs/AWWs Always Ask ANM to Demonstrate How to Use Tools at Subcenter Meeting[a] | 34.2 | 46.4 | 12.1** | 0.006 |
| Topics Discussed at Subcenter Meetings | | | | |
| IFA tablets | 83.2 | 89.7 | 6.5** | 0.032 |
| Arranging transportation | 99.1 | 99.6 | 0.5 | 0.569 |
| Cord care | 52.9 | 69.4 | 16.4** | 0.000 |
| Immediate breastfeeding | 61.8 | 77.5 | 15.7** | 0.000 |
| Feeding child semisolid food | 98.6 | 100.0 | 1.4 | 0.129 |
| Family planning | 97.9 | 100.0 | 2.1* | 0.054 |
| **Working in a Team** | | | | |
| Consider … Part of Their Team | | | | |
| The other FLW of the village (other cadre) | 84.0 | 88.4 | 4.3 | 0.115 |
| The subcenter ANM | 78.9 | 86.7 | 7.8** | 0.038 |
| Other same-cadre FLWs at subcenter | 6.6 | 9.3 | 2.7 | 0.335 |
| Other-cadre FLWs at the subcenter | 8.0 | 11.3 | 3.3 | 0.253 |
| Can Always Get Help from Team When Needed[b] | 59.4 | 62.0 | 2.6 | 0.603 |
| Always Expected to Plan with Team[b] | 52.5 | 68.8 | 16.3** | 0.002 |
| Always Expected to Meet Regularly with Team[b] | 56.1 | 65.3 | 9.2** | 0.049 |

[12] These statistically significant outcomes are strongly related and, if one measure is statistically different across groups, we would expect the others to be as well. Therefore, these differences can be thought of as indicating a single underlying significant difference at baseline, which we would expect given the number of outcomes examined.

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **Working with ANM** | | | | |
| Times Met with ANM Outside Subcenter Meetings in Past Three Months | 4.06 | 4.30 | 0.25 | 0.395 |
| Any Joint Visits with ANM in Past Month | 35.6 | 41.2 | 5.6 | 0.205 |
| ANM Available on Telephone or in Person When Needed Most of the Time[c] | 58.7 | 67.9 | 9.2* | 0.068 |
| ANM Gives Helpful Advice on How to Deal with Certain Cases Most of the Time[c] | 41.3 | 52.8 | 11.5** | 0.031 |
| **Working with ASHA/AWW (opposite cadre)** | | | | |
| Ever Conduct Joint Visits | 69.0 | 78.3 | 9.3** | 0.017 |
| Average Joint Home Visits in Past Week | 1.24 | 1.77 | 0.53** | 0.000 |
| Average Times Met with ASHA/AWW in Past Week to Discuss Work | 1.53 | 2.14 | 0.61** | 0.000 |
| Average Times Asked ASHA/AWW to Conduct Visit (because you could not) in Past 30 Days | 1.01 | 1.43 | 0.42** | 0.011 |
| Average Times ASHA/AWW Asked to Conduct Visit (because she could not) in Past 30 Days | 0.96 | 1.18 | 0.22* | 0.083 |
| **Sample Size** | **273** | **273** | | |
| **ANM Sample** | | | | |
| **Subcenter Meetings** | | | | |
| Attended Three or More Subcenter Meetings in Past Three Months | 80.4 | 92.0 | 11.6* | 0.098 |
| Ran Last Subcenter Meeting Herself | 35.3 | 20.2 | -15.1 | 0.118 |
| Usually Review ASHA/AWW Registers on Own | 31.4 | 13.1 | -18.3** | 0.041 |
| ASHAs/AWWs Always Ask ANM to Demonstrate How to Use Tools at Subcenter Meeting[a] | 37.3 | 47.1 | 9.8 | 0.381 |
| Feel Need for Additional Skills to Lead Meetings | 43.1 | 72.6 | 29.4** | 0.008 |
| **Working in a Team** | | | | |
| ASHAs/AWWs Coordinate Home Visit Efforts Most of the Time[c] | 60.8 | 73.9 | 13.1 | 0.210 |
| Provide ASHA/AWW Advice on Coordination Most of the Time[c] | 70.6 | 76.3 | 5.7 | 0.559 |
| Can Always Get Help from Team When Needed[b] | 56.9 | 73.9 | 17.0 | 0.110 |
| Always Expected to Plan with Team[b] | 49.0 | 68.1 | 19.1* | 0.081 |
| Always Expected to Meet Regularly with Team[b] | 54.9 | 68.7 | 13.8 | 0.184 |
| **Working with ASHAs and AWWs** | | | | |
| Times Per Month Meet with Each ASHA Outside of Subcenter Meetings | 6.60 | 6.98 | 0.38 | 0.670 |
| Times Per Month Meet with Each AWW Outside of Subcenter Meetings | 5.95 | 6.16 | 0.21 | 0.799 |
| Hours Per Week Available to Help ASHAs/AWWs Outside Subcenter Meetings | 3.86 | 3.74 | -0.12 | 0.870 |
| Ever Conduct Joint Visits with ASHA | 84.3 | 95.4 | 11.0* | 0.056 |
| Joint Visits Conducted with ASHA in Past 7 Days | 1.35 | 1.54 | 0.19 | 0.501 |
| Ever Conduct Joint Visits with AWW | 78.4 | 92.7 | 14.2** | 0.034 |
| Joint Visits Conducted with AWW in Past 7 Days | 1.14 | 1.59 | 0.45 | 0.110 |
| **Sample Size** | **51** | **41** | | |

Source: Ananya TBGI RCT Endline.

Notes: Item-specific nonresponse might limit sample size for some comparisons.

**For ASHA/AWW sample:** Regression analysis uses linear probability model and controls for subcenter characteristics (share first births, share Muslim, share SC/ST, and share literate); FLW cadre; stratum-fixed effects; and FLW characteristics (age and indicators for SC/ST status if non-Muslim, Muslim, high education [above 10th standard for ASHA, above 12th for AWW], and the village served being one hour or more travel time from the subcenter). All FLW characteristics and stratum-fixed effects are also interacted with an indicator for FLW cadre.

**For ANM sample:** Regression analysis uses linear probability model and controls for stratum-fixed effects.

[a]Response options are almost never, infrequently, sometimes, and all the time.

[b]Response options are always, sometimes, and never.

[c]Response options are most of the time, some of the time, and none of the time.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

Our analysis shows that FLW attendance at the subcenter meetings was higher in treatment subcenters; whereas 82 percent of control ASHAs and AWWs attended all such meetings in the three months before the survey, 94 percent of treatment ASHAs and AWWs did so, a statistically significant difference. The difference is similar in magnitude and marginally significant when one considers ANM attendance (80 percent of the control group vs. 92 percent of the treatment group). To assess whether there was more active engagement of the FLWs in the treatment areas, we asked all three groups how often ASHAs and AWWs asked ANMs to demonstrate the use of tools (including planners or registers, models, or sample equipment) during the subcenter meetings. According to both ASHA/AWW and ANM reports, ASHAs and AWWs in treatment subcenters show a significantly higher propensity to ask their ANMs about using tools at these meetings (see Figure IV.1).[13]

ASHAs and AWWs further report having more commonly discussed topics directly linked to TBGI targets at subcenter meetings. Discussions of IFA tablets, umbilical cord care, and immediate breastfeeding were significantly more commonly reported at meetings in treatment subcenters, with differences of 7, 16, and 16 percentage points, respectively (compared to mean control group values of 83, 53, and 62 percent). Other differences in discussion topics were insignificant or only marginally significant; however, this is due largely to very high rates of discussion in all subcenters. For example, 100 percent of treatment ASHAs and AWWs were exposed to information on complementary feeding at a subcenter meeting, as were 99 percent of control ASHAs and AWWs. These findings of high reports on some topics are not surprising, because they were part of the basic intervention provided as part of the Ananya program in both the treatment and control subcenters. Subcenter meetings were an important platform for the TBGI intervention and there is some evidence that the intervention changed the nature of these meetings in treatment subcenters. Specifically, ANMs in treatment areas were more likely than those in control areas to report having received assistance from CARE staff during their last subcenter meetings. About 31 percent of control group ANMs reviewed ASHA/AWW registers on their own at this last meeting, compared with only 13 percent of treatment ANMs (a statistically significant reduction of 51 percent). This is likely related to the role of CARE staff in gathering quarterly reports on whether targets were met,

[13] This question was added at the request of a team from Georgia Tech that provided support to CARE in the formative stages of developing the intervention.

**Figure IV.1. An ANM Leading a Discussion at a Subcenter Meeting**

and in helping FLWs to assess progress toward targets more generally. This might have necessitated a more comprehensive review of FLW registers and hence greater involvement of CARE staff compared with the review typically conducted as part of Ananya. Moreover, 35 percent of ANMs at control subcenters ran the most recent subcenter meeting they attended by themselves, compared with only 20 percent of treatment ANMs, which may also be explained partly by the presence of CARE facilitator staff related to the TBGI intervention and their review of information maintained by the FLWs related to their progress on the innovation. This difference was also associated with a greater desire for leadership skills by the ANMs in treatment subcenters. Although only 43 percent of control ANMs felt they should acquire additional skills to better lead subcenter meetings, 73 percent of treatment ANMs responded in this way and commonly cited gaps in teaching or facilitation skills. The difference is large in magnitude (at 68 percent of the control-group mean) and statistically significant. One possible explanation for these findings is that the higher need for setting goals and targets and facilitating meetings was more challenging in the treatment subcenters, so that ANMs were either less confident in their ability to explain these in subcenter meetings; alternatively, their close engagement with CARE staff in these meetings might have highlighted gaps in leadership skills among ANMs.

The TBGI intervention also influenced how the FLWs work together. Those in the treatment subcenters reported that they were more commonly expected to meet and plan with their teams (though not all differences were statistically significant). Additionally, 87 percent of ASHAs and AWWs in treatment subcenters view the ANM at their subcenter as part of their team compared to just 79 percent of those in control areas (a statistically significant difference of 8 percent). ASHAs and AWWs in treatment subcenters reported that the ANM was more often available to assist them and (62 percent said they could always get help when needed, compared to 59 percent of the control group) and more commonly gave helpful advice (53 percent said the ANM gave helpful advice most of the time, compared to 41 percent of control ASHAs and AWWs). Further, both ANMs and ASHAs/AWWs suggested that meetings occurred more frequently between the FLWs. Specifically, ASHAs and AWWs in treatment subcenters conducted significantly more joint visits and met much

more often with one another to discuss work than ASHAs and AWWs in the control subcenters. ASHAs and AWWs in the treatment subcenters also reported asking each other for assistance in conducting home visits significantly more often than their counterparts in control subcenters. For example, 78 percent of ASHAs and AWWs in treatment areas had ever conducted a joint visit, compared to 69 percent of those in control areas. These findings are consistent with the qualitative findings that the TBGI intervention led to increased teamwork between ASHAs and AWWs, including through joint visits, and increased willingness for ASHAs and AWWs to reach out to ANMs for assistance as needed.

## C. Impacts of TBGI on Key Outcomes: Services Provided

Despite the increase in teamwork among FLWs, the TBGI intervention could lead to improved outcomes for beneficiaries only if it led to changes in the extent or types of interactions with households. Our analysis suggests that several, but not all, aspects of service provision improved with the TBGI intervention. Table IV.4 contains impact estimates for the combined sample of ASHAs/AWWs and ANMs.

**Table IV.4. Services Provided by FLWs (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **ASHA/AWW Sample** | | | | |
| **Planning for Home Visits** | | | | |
| Has a Home Visit Planner (observed) | 60.2 | 71.7 | 11.5** | 0.039 |
| Investigator's Perceptions of Planner | | | | |
| Fully updated and in order | 11.5 | 19.6 | 8.0** | 0.013 |
| Somewhat updated and in order | 20.2 | 26.0 | 5.8** | 0.034 |
| Few/no updates | 27.7 | 25.8 | -1.9 | 0.677 |
| **Conducting Home Visits** | | | | |
| Visits to Any Pregnant Women in Last Trimester in Past Month | | | | |
| Average number of women visited (reported by FLW) | 4.06 | 4.40 | 0.35 | 0.154 |
| Number of women visited (verified using planner) | | | | |
| Cannot be verified | 44.5 | 29.8 | -1.5** | 0.000 |
| 0 | 0.2 | 0.0 | -0.3 | 0.318 |
| 1 to 3 | 24.1 | 29.3 | 5.2 | 0.134 |
| 4 or 5 | 21.9 | 23.1 | 1.2 | 0.736 |
| 6 or More | 9.3 | 17.8 | 8.5** | 0.029 |
| Information/services provided to last pregnant woman visited (spontaneous responses) | | | | |
| Provided telephone number for ambulance or private vehicle to take the woman to the hospital for delivery | 65.3 | 88.0 | 22.7** | 0.000 |
| Provided telephone number of ASHA or AWW | 59.0 | 64.7 | 5.7 | 0.166 |
| Told her to apply nothing on the baby's umbilical cord | 80.0 | 98.1 | 18.1** | 0.000 |
| Provided 90 or more IFA tablets | 47.8 | 52.3 | 4.5 | 0.385 |
| Advised her to start breastfeeding with one hour of birth | 100.0 | 99.4 | -0.6 | 0.163 |
| Visits to Any Women Who Delivered (within 24 hours of birth) in the Past Calendar Month | | | | |
| Average number of women visited (reported by FLW) | 2.40 | 2.71 | 0.31 | 0.107 |
| Number of women visited (verified using planner) | | | | |
| Cannot be verified | 46.5 | 38.2 | -8.2* | 0.073 |
| 0 | 0.0 | 0.0 | 0.0 | 1.000 |
| 1 or 2 | 30.6 | 33.3 | 2.6 | 0.546 |
| 3 or 4 | 16.8 | 18.0 | 1.2 | 0.759 |

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| 5 or more | 6.1 | 10.5 | 4.3* | 0.069 |
| Visits to Any Women of Children Ages 5 to 7 Months in Past Calendar Month | | | | |
| Average number of women visited (reported by FLW) | 4.88 | 5.25 | 0.37* | 0.092 |
| Number of women visited (verified using planner) | | | | |
| Cannot be verified | 54.9 | 42.0 | -12.9** | 0.001 |
| 0 | 0.0 | 0.0 | 0.0 | 1.000 |
| 1 or 2 | 6.0 | 4.1 | -1.9 | 0.372 |
| 3 to 5 | 24.3 | 29.3 | 5.1 | 0.178 |
| 6 or more | 14.8 | 24.5 | 9.7** | 0.001 |
| Information provided at last visit | | | | |
| Feed child 2 or 3 *katoris* of food a day | 51.2 | 54.0 | 2.8 | 0.526 |
| Types of food FLW advised woman to feed child (spontaneous responses) | | | | |
| Cereal-based food (rice and daal, roti and milk, or halwa) | 98.6 | 95.8 | -2.8* | 0.091 |
| Fruits or vegetables | 31.5 | 36.6 | 5.1 | 0.256 |
| Meat, fish, or eggs | 9.2 | 13.9 | 4.8** | 0.036 |
| Visits Related to Immunization in Past Calendar Month | | | | |
| Discussed benefits of immunization | 98.4 | 98.5 | 0.1 | 0.905 |
| Talked to women who did not give children follow-up shots | 86.5 | 80.4 | -6.0 | 0.145 |
| Any Discussion of Family Planning During Visits in Past Month | 97.9 | 99.4 | 1.5 | 0.123 |
| **Sample Size** | **273** | **273** | | |
| **ANM Sample** | | | | |
| Share of Time Spent | | | | |
| At subcenter | 40.5 | 37.4 | -3.1 | 0.205 |
| In villages | 33.7 | 34.4 | 0.7 | 0.616 |
| In PHC | 25.8 | 28.2 | 2.4 | 0.277 |
| Performed ANC Checkups in the Past Calendar Month | 98.0 | 97.5 | -0.5 | 0.878 |
| Identified Newborn with Danger Signs and Referred to Appropriate Care in Past Calendar Month | 39.2 | 57.8 | 18.6 | 0.108 |
| Assisted Woman with Serious Postpartum Complications to See Doctor in Past Calendar Month | 70.6 | 89.9 | 19.3** | 0.013 |
| Distributed Contraception in Past Calendar Month | 76.5 | 64.5 | -12.0 | 0.199 |
| **Sample Size** | **51** | **41** | | |

Source: Ananya TBGI RCT Endline.

Notes: Item-specific nonresponse may limit sample size for some comparisons.

**For ASHA/AWW sample:** Regression analysis uses linear probability model and controls for subcenter characteristics (share first births, share Muslim, share SC/ST, and share literate), FLW cadre, stratum-fixed effects, and FLW characteristics (age, and indicators for SC/ST status if non-Muslim, Muslim, high education [above 10th standard for ASHA, above 12th for AWW], and the village served being one hour or more travel time from the subcenter). All FLW characteristics and stratum-fixed effects are also interacted with an indicator for FLW cadre.

**For ANM sample:** Regression analysis uses linear probability model and controls for stratum-fixed effects.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

In preparing and planning for home visits, ASHAs and AWWs at treatment subcenters were far more likely than those at control subcenters to use and maintain a home visit planner. This register, developed by CARE for the Ananya intervention, enables FLWs to build and follow a visit schedule and ensure that all households with pregnant women and children younger than 2 years old receive visits and services at the appropriate times. During our surveys, we asked our interviewers to examine the home visit planner and to assess how complete and filled it was. About 72 percent of ASHAs and AWWs in treatment subcenters produced any home visit planner, compared with only 60 percent of control ASHAs and AWWs, a statistically significant difference of 19 percent. Further, only 32 percent of control FLWs had planners that were fully or somewhat updated and in order, compared with 46 percent of treatment FLWs. Thus, it appears the TBGI intervention has resulted in more active use of planners among FLWs.

We also collected information on the frequency of home visits, in terms of the number of women visited in the previous month. We looked specifically at three types of visits: in the last trimester of pregnancy, immediately after birth, and five to seven months after birth (when children should start eating semisolid foods). In addition, we analyzed both the number of home visits an FLW reported having made and the number of home visits our investigators were able to verify that she made based on her planner. Compared with those in control subcenters, ASHAs and AWWs in treatment subcenters reported conducting on average 0.35 additional visits to pregnant women, (4.40 vs. 4.06 visits), 0.31 additional visits to mothers within a day of birth (2.71 vs. 2.40 visits), and 0.37 additional visits to mothers of 5- to 7-month-old children (5.25 vs. 4.88 visits) in the month before our survey. However, none of the differences were statistically significant at the 5 percent level (though the last was significant at the 10 percent level). FLWs in treatment areas, however, showed more investigator-verified visits to all three groups of women. These differences were likely driven largely by differences in record-keeping—our investigators could verify FLW reports far more often in treatment areas than in control areas and, if FLWs with completed planners are more likely to make visits, then the difference in record-keeping could be driving the difference in investigator-verified visits.

Our study also sought to collect information on whether ASHAs and AWWs were providing information relevant to the TBGI goals during their home visits. Again, we focused on visits to pregnant women in their last trimester, households with newborns, and households with 5- to 7-month-olds. To reduce recall error, we asked the respondent to focus on the most recent visit she made to each type of household. For example, we asked the respondent to think about the most recent visit she made to a pregnant woman in her last trimester of pregnancy and then tell us about the information provided to that woman during any of the home visits conducted during her pregnancy.

Our analysis showed that during household visits, ASHAs and AWWs from treatment subcenters were more likely than those from control subcenters to report having discussed some, but not all, topics related to the TBGI targets. Significantly more treatment FLWs reported providing the telephone number for an ambulance or private vehicle to take the woman to the hospital for delivery (88 versus 65 percent, a difference of 35 percent) and information on clean-cord care (98 versus 80 percent, a difference of 23 percent) during visits to pregnant women. However there was no significant difference in provision of an FLW's telephone number or discussions of IFA tablets or immediate breastfeeding at such visits. The information FLWs presented during visits to women with a child about 6 months old also differed little across the treatment and control groups. Finally, the two sets of ASHAs/AWWs did not exhibit significantly different propensities to discuss family planning or immunizations with households (though reported coverage of these topics was above 98 percent or more in both treatment and control

groups, which made it difficult for the TBGI intervention to lead to large changes). These findings are largely consistent with the fact that, even in the control areas, FLWs should be providing similar messages to the households.

Overall, it appears that the TBGI intervention led to large changes in the ways FLWs coordinated—consistent with the qualitative findings—and smaller, though nontrivial, changes in their delivery of services. To assess whether these changes were associated with improvements in maternal and child health, we turn in the next chapter to the beneficiary survey where we look at beneficiaries' perceptions of the FLWs and their health behaviors and practices.

# V. RESULTS FROM THE RCT: BENEFICIARY SURVEY

In assessing the impact of the TBGI intervention on beneficiaries, we kept the seven TBGI goals in mind. Given the focus of the TBGI incentives on achieving these goals, estimating impacts on outcomes directly related to them was an important focus of the evaluation. However, our impact analysis focused on a broader set of outcomes, including those both directly related and not directly related to the TBGI goals, to provide a comprehensive assessment of the impacts of TBGI. Specifically, because the TBGI intervention incentivized FLW interactions with households around certain behaviors, there might have been spillover effects on nontargeted outcomes. These spillover effects could have been either positive (if FLWs provided additional services or information during goal-focused visits) or negative (if the incentives led to a focus on certain behaviors at the expense of others). Therefore, although we were agnostic about the direction of the effects on outcomes not directly related to the goals, we included them in the impact analysis.

In this chapter, we describe the results for key outcomes from the beneficiary survey. We begin by assessing whether the random assignment design was successful in creating equivalent treatment and control groups of beneficiaries at baseline and endline. Next, we examine impacts on FLW interactions with beneficiaries from the beneficiaries' perspective, and beneficiaries' knowledge that might have been expected to mediate changes in behavior. We then assess impacts on outcomes directly related to the seven TBGI goals and explore impacts on a broader set of outcomes not directly targeted by TBGI in each of the following domains: antenatal care and delivery preparation, delivery and postpartum care, child nutrition, child immunizations, and reproductive health. Finally, we examine the extent to which impacts varied across key subgroups of interest. Because we examine impacts on a large set of outcomes, we have to be mindful that some differences can be significant simply by chance (this is known as the multiple comparisons issue). We therefore take care to examine the overall pattern of impacts in each domain before drawing conclusions about the impacts of the intervention.

## A. Baseline Equivalence

Random assignment should ensure that the treatment and control groups are, on average, statistically equivalent at baseline. However, an unlucky randomization draw can cause treatment and control groups to differ by chance. Therefore, to verify the similarity of the two groups of beneficiary households, we used our baseline data to compare demographic characteristics and key outcomes across the domains targeted by the survey (Table V.1). These comparisons suggest that treatment and control groups were statistically very similar. Only one of the 22 baseline differences we considered (use of permanent methods of contraception) was statistically significant—no more

**Table V.1. Baseline Differences Between Treatment and Control (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Demographic Characteristics of Respondent | | | | |
| Hindu | 88.9 | 89.4 | 0.5 | 0.921 |
| Scheduled caste/scheduled tribe (SC/ST) | 30.6 | 35.9 | 5.3 | 0.414 |
| Age (years) | 25.2 | 25.5 | 0.3 | 0.386 |
| Birth parity (mean) | 2.38 | 2.36 | -0.02 | 0.853 |
| Illiterate | 63.0 | 58.1 | -4.9 | 0.241 |
| SES quartile (mean)[a] | 2.78 | 2.81 | 0.03 | 0.834 |
| Antenatal Care | | | | |
| At least 3 ANC visits | 26.3 | 26.0 | -0.3 | 0.941 |
| Received 90 IFA tablets | 25.2 | 22.4 | -2.9 | 0.552 |
| Consumed 90 IFA tablets | 13.0 | 11.3 | -1.7 | 0.660 |
| Delivery and Postpartum Care | | | | |
| Facility delivery | 69.2 | 71.3 | 2.2 | 0.609 |
| Nothing applied to cord | 37.7 | 42.8 | 5.1 | 0.501 |
| Immediate breastfeeding | 50.0 | 49.5 | -0.5 | 0.919 |
| Exclusive breastfeeding for 6 months, children 6 months or older[b] | 31.1 | 37.5 | 6.3 | 0.407 |
| Exclusive breastfeeding in past 24 hours, children younger than 6 months[c] | 52.7 | 45.5 | -7.2 | 0.312 |
| Child Nutrition, Children 6 Months or Older | | | | |
| Child eats solid or semisolid food | 47.5 | 50.5 | 3.0 | 0.580 |
| Child ate cereal-based meal in previous day | 41.1 | 45.0 | 3.8 | 0.453 |
| Immunizations, Children 6 Months or Older | | | | |
| Child received DPT3 | 49.4 | 56.7 | 7.4 | 0.227 |
| Child fully immunized (except measles) | 39.9 | 45.5 | 5.6 | 0.370 |
| Reproductive Health, All Women | | | | |
| Use of permanent methods | 11.6 | 6.8 | -4.8 | 0.027** |
| Use of any modern method | 21.4 | 18.6 | -2.7 | 0.505 |
| FLW Interactions | | | | |
| FLW visit in final trimester | 45.1 | 42.3 | -2.9 | 0.634 |
| FLW visit in first month after delivery | 12.1 | 13.6 | 1.5 | 0.739 |

Source: Ananya TBGI RCT Baseline.

Notes: Treatment means and treatment–control differences are adjusted for differences in random assignment stratum using an ordinary least squares regression with stratum-fixed effects. Reported *p*-values account for clustering of standard errors at the subcenter level. Sample is restricted to the 36 treatment and 36 control subcenters surveyed at endline.

Sample sizes are 1,574 to 1,596 (all women), 804 (children younger than 6 months), and 768 (children 6 months or older).

[a]SES quartile determined using coefficients and cutoffs from a principal components analysis using the Ananya statewide baseline data (following the methodology of the National Family Health Survey's wealth index). Quartiles are therefore relative to the statewide SES distribution for women who gave birth in the previous 12 months.

[b]Based on self-reports of the duration of exclusive breastfeeding for children 6 months or older.

[c]Based on reports of liquids and solids fed to children younger than 6 months old in the previous 24 hours, following the recommended definition of the World Health Organization.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

than would be expected by chance—though a handful of the differences were relatively large in magnitude. These results suggest that the random assignment was successful in creating equivalent treatment and control groups and increase our confidence that any statistically significant differences between the groups at endline are unlikely to be spurious. Nevertheless, to account for some of the baseline differences that we did observe, we controlled for baseline levels of the outcome aggregated to the subcenter level in the impact analysis.

## B. Endline Demographic and Socioeconomic Characteristics

The demographic and socioeconomic characteristics of the surveyed beneficiaries provide important context about the population targeted by the TBGI intervention. In addition, because the demographic and socioeconomic characteristics of these respondents might be correlated with the outcomes of interest, it is important to confirm that these characteristics are similar between the treatment and control groups at endline. Otherwise, chance differences in the characteristics (which are unlikely to have been affected by the intervention) might be driving the observed impacts. We observe that the sample of beneficiaries was quite disadvantaged (Table V.2)—about a third of the treatment and control groups belonged to scheduled castes or tribes, about half had no formal education whatsoever, and a similar fraction was illiterate. However, compared with the overall population of women who had recently given birth in Bihar, the sample was slightly better off in terms of SES, with about 60 percent of women in the top two statewide SES quartiles (compared with 50 percent across the state, by definition). The treatment and control groups were very similar in all the demographic characteristics that we considered—the differences are mostly small in magnitude and none are statistically significant. This suggests that differences in demographic and socioeconomic characteristics likely did not drive any impacts that we observed, though we did control for these small differences in the impact analysis to improve the precision of the estimates.

## C. Impacts on FLW Interactions

To complement our earlier analysis of FLW–household interactions as reported by FLWs, we examined the extent to which the TBGI intervention affected the quantity and nature of these interactions, as perceived by the beneficiaries themselves (Table V.3). We observe large and statistically significant impacts across almost all the outcomes relating to home visits and FLW advice. In the absence of the intervention, 73 percent of the control group reported an FLW visit during the final trimester of pregnancy and 38 percent reported a visit in the first 24 hours after delivery; in the treatment group, these rates were both about 7 percentage points higher (10 and 18 percent of the respective control means). These findings are largely consistent with those from the FLW surveys that showed an increase in the number of home visits, although the FLW findings were not always statistically significant and are not directly comparable because they focus on the number of home visits rather than the rate. In addition, beneficiaries in the treatment group were significantly more likely than control group members to have received advice on key topics, both those related directly to the TBGI goals (for example, regarding IFA tablets and transportation) and other topics (for example, advice on danger signs, saving money for delivery, and identifying a facility for delivery). Although the probability of an FLW visit in the first month after delivery was not significantly different between the treatment and control groups, beneficiaries were more likely to have received advice on key topics, such as infant danger signs and exclusive breastfeeding for infants.

**Table V.2. Endline Demographic and Socioeconomic Characteristics of Respondents (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value[b] |
|---|---|---|---|---|
| Hindu | 88.6 | 91.0 | 2.4 | 0.596 |
| SC/ST | 31.9 | 34.4 | 2.5 | 0.650 |
| Household Size (number of people) | 6.1 | 6.2 | 0.1 | 0.476 |
| Age | | | | |
| 15–19 | 7.5 | 6.3 | -1.2 | 0.254 |
| 20–24 | 42.3 | 41.5 | -0.8 | 0.786 |
| 25–29 | 31.7 | 36.4 | 4.7 | 0.051* |
| 30–34 | 13.3 | 10.5 | -2.8 | 0.102 |
| 35–49 | 5.2 | 5.3 | 0.1 | 0.945 |
| | | | | 0.209[u] |
| Mean (years) | 24.9 | 25.0 | 0.1 | 0.819 |
| Birth Parity | | | | |
| 1 child | 29.5 | 30.6 | 1.2 | 0.546 |
| 2 children | 25.8 | 27.3 | 1.5 | 0.488 |
| 3 children | 21.6 | 21.2 | -0.3 | 0.881 |
| 4 or more children | 23.2 | 20.8 | -2.3 | 0.292 |
| | | | | 0.684[u] |
| Mean (number of children) | 2.60 | 2.53 | -0.07 | 0.310 |
| No Formal Education | 47.0 | 49.8 | 2.8 | 0.390 |
| Illiterate | 49.7 | 53.1 | 3.4 | 0.335 |
| BPL Card | 55.1 | 51.4 | -3.7 | 0.312 |
| SES Quartile[a] | | | | |
| Quartile 1 | 17.4 | 22.2 | 4.9 | 0.141 |
| Quartile 2 | 21.0 | 21.1 | 0.1 | 0.966 |
| Quartile 3 | 34.8 | 32.1 | -2.7 | 0.387 |
| Quartile 4 | 26.8 | 24.6 | -2.3 | 0.613 |
| | | | | 0.512[u] |
| Mean (quartile) | 2.71 | 2.59 | -0.12 | 0.292 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted for differences by random assignment stratum using a regression with stratum-fixed effects. Reported *p*-values account for clustering of standard errors at the subcenter level. Tests for differences in distributions (age, birth parity, and SES quartile) were conducted using seemingly unrelated estimation regression in Stata. Sample includes 36 treatment and 36 control subcenters.

Sample size is 1,598 to 1,607.

[a]SES quartile determined using coefficients and cutoffs from a principal components analysis using the Ananya statewide baseline data (following the methodology of the National Family Health Survey's wealth index). Quartiles are therefore relative to the statewide SES distribution for women who gave birth in the previous 12 months.

[b]*p*-values are for the test of equivalence of distributions.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

**Table V.3. Impacts on FLW Interactions: Advice and Home Visits (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **Pregnancy and Delivery** | | | | |
| Visits in Last Trimester | | | | |
| ASHA or AWW visited in last trimester | 73.1 | 80.5 | 7.4 | 0.006** |
| Number of visits in last trimester | 1.57 | 1.98 | 0.41 | 0.000** |
| Advice Received | | | | |
| FLW Gave Advice on TT injections | 87.1 | 84.0 | -3.1 | 0.176 |
| Advice on IFA Tablets | | | | |
| FLW gave advice | 60.5 | 67.7 | 7.3 | 0.036** |
| FLW told woman to consume at least 90 IFA tablets | 24.2 | 37.5 | 13.3 | 0.000** |
| FLW explained benefits | 30.9 | 37.7 | 6.8 | 0.045** |
| Advice on Saving Money for Delivery | 44.3 | 57.8 | 13.5 | 0.000** |
| Advice on Identifying Facility for Delivery | 33.1 | 41.6 | 8.5 | 0.014** |
| Advice on Any Maternal Danger Signs | 32.9 | 43.7 | 10.8 | 0.004** |
| Information on Transportation | | | | |
| ASHA or AWW gave number of ambulance | 12.9 | 19.1 | 6.2 | 0.023** |
| ASHA or AWW gave number of private vehicle | 18.2 | 26.4 | 8.2 | 0.004** |
| ASHA or AWW gave own number | 51.5 | 56.2 | 4.6 | 0.219 |
| ASHA or AWW Visited in First 24 Hours After Delivery | 37.9 | 44.8 | 7.0 | 0.018** |
| **Postpartum Care** | | | | |
| ASHA or AWW Visited After 24 Hours but Within First Month After Delivery | 53.6 | 57.7 | 4.1 | 0.367 |
| Number of Visits in First Month After Delivery | 1.25 | 1.50 | 0.24 | 0.063* |
| Advice by FLW in First Month After Delivery | | | | |
| Any advice on infant danger signs | 25.6 | 35.2 | 9.6 | 0.004** |
| Advice on exclusive breastfeeding for 6 months | 49.5 | 56.7 | 7.2 | 0.098* |
| **Child Nutrition** | | | | |
| FLW Visits Related to Complementary Feeding (children 5 months or older) | | | | |
| Any visit related to complementary feeding | 24.9 | 40.9 | 16.0 | 0.000** |
| Advised to start feeding at age 6 months | 16.3 | 25.5 | 9.2 | 0.024** |
| Advice on types of food | 10.5 | 25.0 | 14.5 | 0.000** |
| Advice on times to feed | 12.7 | 26.7 | 14.0 | 0.000** |
| Advised on quantity of food using *katori* | 10.2 | 25.1 | 14.9 | 0.000** |
| Advised to feed from separate bowl | 11.3 | 22.9 | 11.6 | 0.002** |
| **Reproductive Health** | | | | |
| FLW Visit to Discuss Family Planning | 18.2 | 28.4 | 10.2 | 0.000** |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic),;and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters.

Sample sizes are 1,607 (all women) and 939 (children 5 months or older).

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

Impacts on FLW visits around complementary feeding for children 5 months or older were uniformly large, positive, and strongly significant. Although only 25 percent of beneficiaries in the control group received such a visit, about 41 percent of beneficiaries in the treatment group received one (a difference of 16 percentage points, or 64 percent of the control mean). Similarly, beneficiaries in the treatment group were far more likely to have received advice on issues such as the initiation of complementary feeding at age 6 months, the type of food, and the frequency and quantity of feeding—all areas emphasized by the TBGI goal around child nutrition. Similarly, 28 percent of the treatment group reported a visit by an FLW to discuss family planning compared with 18 percent of the control group (an impact of about 10 percentage points, or 56 percent of the control mean), which relates to the TBGI goal promoting use of modern contraceptive methods. These findings are consistent with the qualitative findings that FLWs used the TBGI goals to focus the messages that they delivered during home visits. Although the impact on discussion of specific topics by FLWs in their most recent visit with households as reported in the FLW surveys was more variable, this is partly due to the fact that FLWs reported much higher rates of discussion—limiting the scope for impact. The higher reported rates by FLWs could reflect, for example, that they gave the socially desirable response or that they did not reach all relevant women so that their most recent visit was not representative of the experience of the average woman.

In terms of the features of home visits, beneficiaries in the treatment group were significantly more likely than control group members to report that an FLW ever used the various job-aid tools (such as mobile *Kunji* cards, which are illustrated plastic cards with key health messages) included as part of the Ananya program to help promote behavior change (see Figure V.1 and Table V.4). Finally, because the structure of the TBGI incentives aimed to promote cooperation among various cadres of FLWs serving the same catchment area, we examined impacts on joint visits. We found statistically significant large positive impacts on the probability of a respondent reporting a joint visit by an ASHA and AWW. In the treatment group, 37 percent of beneficiaries reported such

**Figure V.1. Job Aids Used During Home Visits**

Mobile Kunji

Katori and Spoon

Uterus Model


Mala-D Contraceptive Pills

**Table V.4. Impacts on FLW Interactions: Features of Home Visits (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Duration of Most Recent FLW Visit | | | | |
| No visit (0 minutes) | 21.6 | 14.6 | -7.0 | 0.005** |
| Fewer than 5 minutes | 32.9 | 29.9 | -3.1 | 0.337 |
| 6–15 minutes | 31.6 | 37.5 | 5.9 | 0.119 |
| 16–30 minutes | 12.2 | 14.5 | 2.2 | 0.346 |
| More than 30 minutes | 1.3 | 3.5 | 2.2 | 0.032** 0.027**[a] |
| Average, among those with visits (minutes) | 10.7 | 12.3 | 1.6 | 0.050** |
| Talked to Husband in Most Recent Visit | 14.6 | 20.0 | 5.4 | 0.124 |
| Talked to Mother-in-Law in Most Recent Visit | 11.2 | 15.6 | 4.3 | 0.045** |
| Job Tools Ever Used by FLW | | | | |
| Mobile Kunji cards | 18.8 | 28.8 | 10.0 | 0.004** |
| Dr. Anita | 16.9 | 26.6 | 9.6 | 0.002** |
| Katori/spoon | 10.7 | 21.5 | 10.7 | 0.000** |
| Uterus model | 5.5 | 13.6 | 8.0 | 0.003** |
| Copper-T IUD | 5.5 | 12.4 | 6.9 | 0.001** |
| Mala-D contraceptive pills | 6.5 | 12.9 | 6.3 | 0.003** |
| ASHA and AWW Ever Visited Together | 27.2 | 37.0 | 9.7 | 0.003** |
| ANM and ASHA/AWW Ever Visited Together | 22.6 | 32.0 | 9.4 | 0.010** |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Tests for differences in distributions (duration of visit) were conducted using seemingly unrelated estimation regression in Stata. Sample includes 36 treatment and 36 control subcenters.

Sample sizes are 1,607 (all women) and 1,321 (women with visits).

[a] *p*-value is for the test of equivalence of distributions.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

cadres of FLWs serving the same catchment area, we examined impacts on joint visits. We found statistically significant large positive impacts on the probability of a respondent reporting a joint visit by an ASHA and AWW. In the treatment group, 37 percent of beneficiaries reported such a visit versus 27 percent of the control group (a difference of 10 percentage points, or 36 percent of the control mean). Similarly, about 32 percent of the treatment group reported a joint visit by an ANM and an ASHA/AWW versus 23 percent of the control group (a difference of 9 percentage points, or 42 percent of the control mean). This increase in joint visits is consistent with both the qualitative findings and the findings from the FLW surveys.

## D. Impacts on Mothers' Knowledge

An important mechanism through which improved interactions with FLWs could affect maternal and child health-related behaviors is by increasing beneficiaries' knowledge of issues related to their health and the health of their child, and of desirable health behaviors. We therefore assessed impacts on beneficiaries' knowledge of these issues and behaviors in several domains (Table V.5).

**Table V.5. Impacts on Mother's Knowledge (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Knowledge of Maternal Danger Signs During Pregnancy and Delivery | | | | |
| Prolonged labor | 29.1 | 26.8 | -2.4 | 0.444 |
| Excessive bleeding | 45.6 | 43.0 | -2.7 | 0.518 |
| Convulsions | 21.4 | 20.8 | -0.7 | 0.832 |
| Swelling of hands, body, or face | 19.0 | 25.7 | 6.8 | 0.011** |
| Fever | 27.2 | 27.9 | 0.7 | 0.842 |
| Vaginal discharge | 3.7 | 3.2 | -0.5 | 0.755 |
| Severe abdominal pain | 6.3 | 7.2 | 0.9 | 0.554 |
| Knowledge of Maternal Danger Signs in First 6 Weeks | | | | |
| Excessive bleeding | 33.5 | 36.9 | 3.4 | 0.403 |
| Severe abdominal pain | 42.6 | 32.3 | -10.2 | 0.018** |
| Fever | 50.4 | 48.3 | -2.1 | 0.573 |
| Vaginal discharge | 8.7 | 12.1 | 3.3 | 0.108 |
| Severe headache or blurred vision | 6.4 | 9.1 | 2.7 | 0.144 |
| Convulsions | 8.3 | 7.8 | -0.5 | 0.794 |
| Fits | 5.3 | 6.9 | 1.6 | 0.362 |
| Knowledge of Infant Danger Signs | | | | |
| Diarrhea | 33.1 | 35.6 | 2.6 | 0.469 |
| Fever | 61.0 | 59.3 | -1.7 | 0.716 |
| Cough/cold | 57.9 | 51.8 | -6.1 | 0.182 |
| Breathing difficulties | 8.9 | 9.7 | 0.9 | 0.674 |
| Infant not crying | 4.9 | 6.0 | 1.1 | 0.523 |
| Chest problems | 4.0 | 4.7 | 0.7 | 0.659 |
| Blue tongue and lips | 1.3 | 1.6 | 0.3 | 0.740 |
| Baby not taking milk | 3.0 | 3.0 | -0.1 | 0.948 |
| Pneumonia | 16.5 | 11.3 | -5.2 | 0.058* |
| Baby not gaining weight | 2.1 | 2.7 | 0.6 | 0.581 |
| Baby small or premature | 2.2 | 3.7 | 1.5 | 0.205 |
| Baby is drowsy | 1.0 | 0.0 | -1.1 | 0.021** |
| Baby cold to touch | 4.2 | 4.4 | 0.3 | 0.802 |
| Jaundice | 9.4 | 7.6 | -1.8 | 0.338 |
| Nothing Should Be Applied to Cord | 11.2 | 11.7 | 0.5 | 0.830 |
| Bath Should Be Delayed by at Least 2 Days | 15.7 | 17.3 | 1.6 | 0.590 |
| Should Breastfeed Immediately | 41.4 | 48.7 | 7.4 | 0.008** |
| Should Exclusively Breastfeed for 6 Months | 75.6 | 83.3 | 7.7 | 0.004** |
| Solid Foods Should Be Given Starting at Age 6 Months | 46.3 | 47.0 | 0.7 | 0.902 |
| Knows 3 or More Modern Methods of Contraception | 96.7 | 97.0 | 0.2 | 0.885 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters.

Sample size is 1,577 to 1,607.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

Impacts on knowledge most closely related to the TBGI goals, including knowledge around cord care, starting complementary feeding by age 6 months, and modern methods of contraception, are small and not statistically significant (though knowledge of modern methods was very high overall, which suggests limited scope for improvement). However, there are large and significant impacts on knowledge around immediate breastfeeding and exclusive breastfeeding for 6 months. Knowledge around immediate breastfeeding was 49 percent in the treatment group versus 41 percent in the control group (a difference of about 7 percentage points, or 18 percent of the control mean); knowledge around exclusive breastfeeding was 83 percent in the treatment group versus 76 percent in the control group (a difference of about 8 percentage points, or 10 percent of the control mean). As we show in the next section, these impacts on knowledge, however, did not translate into significant impacts on immediate breastfeeding, but there is some evidence that they are associated with impacts on exclusive breastfeeding.

## E. Impacts on the Targeted Behaviors

The TBGI intervention explicitly focuses on two goals in the antenatal care and delivery preparation domain (making appropriate transportation plans for delivery and receipt of at least 90 IFA tablets) and two goals in the delivery and postpartum care domain (applying nothing to the umbilical cord after cutting and immediate breastfeeding). For practices in these domains to have been affected by TBGI, beneficiaries would have to have been exposed to the intervention during the final trimester of pregnancy or the immediate postpartum period.14 Because the intervention began 12 months before the survey, women who gave birth more than nine completed months ago would not have had full exposure during their pregnancy; we therefore restricted the sample for the analyses in these domains to exclude these women. The impact estimates suggest that there were very limited impacts on the TBGI outcomes in either these domains (Table V.6): all of the impacts are small in magnitude and none are statistically significant.[15]

Child nutrition is another important focus of the TBGI intervention, which explicitly incentivizes the promotion of age-appropriate complementary feeding for children 6 months or older, focusing on the appropriate types of foods and the frequency and quantity of meals. The impact of the TBGI intervention on the percentage of children older than 6 months who are eating solid or semisolid food is positive, about 5 percentage points, but is not statistically significant; neither is the indicator for initiating feeding at age 6 months (Table V.6). However, although the percentage of children being fed complementary foods was similar, there is some evidence that specific feeding practices improved. Specifically, 55 percent of children in the treatment group were fed a cereal-based meal in the previous day versus 46 percent of the control group (a difference of 9 percentage points, or 19 percent of the control mean); 41 percent of the treatment group were fed

[14] More generally, impacts could have been larger for women exposed to the intervention more recently, because there would have been an initial settling in period as the FLWs became accustomed to the incentives. To check this, we conducted the analyses of key outcomes restricted to women who gave birth in the previous six months (pregnancy, delivery, and postpartum care outcomes) or the past nine months (all other outcomes). The overall pattern of impacts with these restrictions (Table V.11) was similar to the full sample, though some impacts were slightly higher and others slightly lower than before.

[15] Because the scope for impacts on immediate breastfeeding might have been limited for facility deliveries, in which this is more a function of the maternal care provided at the facility rather than the woman's decision, we also examined impacts on this outcome restricted to home deliveries. These estimates were similarly small in magnitude and not statistically significant.

from a separate bowl versus 32 percent of the control group (also a difference of 9 percentage points, or 29 percent of the control mean). There were also marginally significant, though small, impacts on the number of times fed (0.2 times, or 16 percent of the control mean of 1.2 times) and the amount fed (0.1 of a *katori,* or 31 percent of the control mean of 0.35 *katoris*) in the previous day.

**Table V.6. Impacts on Outcomes Directly Related to TBGI Goals (percentages unless otherwise indicated)**

| TBGI Goal | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **1. Transportation Plans for Delivery**[a] | | | | |
| Obtained correct number of ambulance | 8.6 | 12.4 | 3.7 | 0.120 |
| Obtained number of private vehicle | 11.2 | 10.8 | -0.5 | 0.851 |
| Obtained number of FLW | 47.1 | 47.6 | 0.4 | 0.919 |
| Obtained any number | 55.4 | 54.5 | -0.8 | 0.829 |
| **2. IFA Tablets**[a] | | | | |
| At least 90 tablets received | 16.0 | 18.0 | 2.0 | 0.516 |
| **3. Early Initiation of Breastfeeding**[a] | | | | |
| Immediate breastfeeding | 56.4 | 58.4 | 2.0 | 0.613 |
| **4. Appropriate Umbilical Cord Care**[a] | | | | |
| Nothing applied to cord | 56.0 | 55.6 | -0.4 | 0.928 |
| Clean-cord care[b] | 47.5 | 48.9 | 1.4 | 0.765 |
| **5. Appropriate Complementary Feeding (ages 6 to 11 months)**[c] | | | | |
| Child eats solid or semisolid food | 62.4 | 67.4 | 5.0 | 0.211 |
| Child began eating solid food by age 6 months | 25.4 | 28.7 | 3.3 | 0.427 |
| Child feeding in previous day | | | | |
| Times fed | 1.24 | 1.45 | 0.21 | 0.085* |
| Fed any meal from separate bowl | 31.7 | 40.9 | 9.2 | 0.047** |
| Fed any cereal-based meal | 46.3 | 55.3 | 9.0 | 0.016** |
| Amount fed (*katoris*) | 0.35 | 0.47 | 0.11 | 0.097* |
| **6. Family Planning Within 6 Months of Delivery**[d] | | | | |
| Use of permanent methods | 7.6 | 4.6 | -3.0 | 0.073* |
| Use of any modern method | 10.5 | 10.3 | -0.2 | 0.957 |
| **7. Children Who Received a DPT3 Injection by Age 6 Months**[c] | | | | |
| Received DPT3 (card and self-reports) | 67.4 | 72.8 | 5.4 | 0.243 |
| Received DPT3 (card only) | 68.5 | 75.8 | 7.3 | 0.113 |
| Received DPT3 (self-reports only) | 61.3 | 65.8 | 4.5 | 0.584 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters.

Sample sizes are 1,046 to 1,200 for goals 1–4; 814 to 822 for goal 5; 767 for goal 6; and 261 to 715 for goal 7.

[a]Sample is women who gave birth in the previous 9 months.

[b]New blade and new thread used, and nothing applied to the cord.

[c]Sample is children ages 6 to 11 months.

[d]Sample is women who gave birth in the previous 6 months.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

Another of the TBGI goals involves child immunization. Because immunization rates are typically high for early routine immunizations but drop off for later ones, the goal focuses on DPT3—one of the later immunizations.[16] To examine the impact of the TBGI intervention on DPT3 immunization (and immunizations more generally), we relied on information from both child immunization cards and, when the card was not available or blank, from mothers' recall. Although the routine immunizations (except for measles) are supposed to be completed by 14 weeks of age, our analysis focused on a slightly older sample of children older than 6 months to allow for some delay in immunizations. Using the combined data from cards and self-reports (Table V.6), we found nearly 73 percent of treatment group children having DPT3 compared to about 67 percent of control group children, although this difference was not statistically significant. Using information from immunization cards only—which is likely to be more reliable than mother's recall—the impact on DPT3 was larger in magnitude (7.3 percentage points) than that estimated using the combined reports (5.4 percentage points); however, this difference was still not statistically significant.

The final TBGI goal, in the reproductive health domain, involved the adoption of modern methods of contraception (either permanent or temporary methods) within six months of giving birth. For the sample of women who gave birth in the previous six months—which make up about half our total endline sample—there is little evidence of positive impacts on the use of modern methods overall (use of permanent methods is significantly different at the 10 percent level, but the difference is negative).

These findings of very modest impacts of the TBGI intervention on the seven goals are in contrast to the self-reported achievements of the treatment subcenters that determined the award of incentives. Over the course of this study, 30 of the 38 subcenters met their quarterly target of five of seven goals and received incentives in all quarters based on reports received from CARE. The remaining 8 subcenters received incentives in all quarters but the first, when FLWs could have still been familiarizing themselves with the intervention. In terms of specific goals, these reports suggest that all goals except for the goal around IFA tablets were met at least 80 percent of the time (and more than 90 percent of the time for the immediate breastfeeding, cord care, and complementary feeding goals). The goal around IFA tablets was met about 70 percent of the time, possibly reflecting problems with the supply of these tablets. It is difficult to compare these reported achievements by subcenters directly with the reports in our beneficiary survey, because we do not have information on the exact coverage percentages reported by subcenters. Nevertheless, the coverage rates of most of the goal-related outcomes in the treatment group in our beneficiary survey were well below the rewarded coverage rates of 70 to 80 percent (or 30 percent for family planning).

Although we cannot be certain of the reasons for this discrepancy, there are several possibilities. First, there might be differences in how the FLWs measure goal attainment, and how we capture it in our surveys. As we found from the FLW reports in our surveys and in our qualitative work, FLWs had a varied understanding of what attaining a goal meant, which ranged from informing households about a practice for some to ensuring the practice was followed. Second, FLWs collected data on targets on a regular basis, whereas we relied on beneficiaries' longer-term recall—difficulties with recall at the time of the follow-up survey might have led us to report lower overall coverage rates. Third, our data included a representative sample of beneficiaries in each subcenter's catchment area.

[16] The full set of routine immunizations includes BCG at birth, OPV1 (polio 1) and DPT1 at age 6 weeks, OPV2 (polio 2) and DPT2 at age 10 weeks, OPV3 (polio 3) and DPT3 at age 14 weeks, and measles at age 9 months.

If some of these households were not included in FLW registers and these left-out households had poorer outcomes (for example, because they came from disadvantaged populations), then the difference in populations could explain these results. However, we expect that CARE efforts to identify left-out households and have FLWs serve them suggests that this seems unlikely to be driving the differences we observe. Finally, FLWs could also be overstating their self-reported coverage rates to receive incentives, and there may also be some type of social desirability bias. Nevertheless, the beneficiary survey results do provide strong evidence that, regardless of whether the targeted coverage rates are truly achieved, the incentives motivate FLWs to exert greater effort and cooperation in their interactions with beneficiaries that could plausibly affect these outcomes.

## F. Impacts on Nonincentivized Behaviors

In addition to outcomes directly related to the two TBGI goals in the antenatal care domain, we also examined impacts on outcomes in this domain that were not directly incentivized (Table V.7), such as specific delivery preparations (for example, saving money for delivery or identifying a facility for emergencies). Similar to our findings for the TBGI outcomes, these impacts were generally small in magnitude and none were statistically significant. Similarly, there was no evidence of significant impacts on most nonincentivized outcomes in the delivery and postpartum care domain (Table V.8), including facility delivery, thermal care, treatment-seeking behavior, and FLW advice for danger signs.[17] The one outcome that had a significant impact was exclusive breastfeeding, measured using the World Health Organization definition of breastfeeding in the past 24 hours for children younger than 6 months (WHO 2010).[18] In the treatment group, 71 percent of beneficiaries were exclusively breastfeeding using this measure, versus 61 percent in the control group (a difference of about 9 percentage points, or 15 percent of the control mean). Given the large number of outcomes considered in this domain, we cannot rule out that this outcome is significant by chance.

However, the alternative breastfeeding measure based on self-reports of the duration of exclusive breastfeeding for children older than 6 months (which is much lower, only 26 percent in the control group) also shows a large positive impact of 8 percentage points, or 30 percent of the control mean, although this is not statistically significant.[19] Therefore, there is some suggestive evidence that exclusive breastfeeding practices did improve even though this was not explicitly targeted by the TBGI intervention.

[17] The impacts on treatment-seeking behavior and FLW advice for danger signs should be viewed with some caution, because they are conditional on experiencing danger signs and any impact of the intervention on identification of danger signs could lead to compositional differences between the treatment and control groups.

[18] This measure avoids the errors associated with the self-reported duration of exdusive breastfeeding, though it reflects only the current exdusive breastfeeding status of young children and not the duration of exdusive breastfeeding.

[19] The restriction to women who gave birth in the previous 9 months (for full exposure to the intervention) and to children who are at least 6 months old (for full exposure to the exdusive breastfeeding period) implies that the outcome reported here applies only to children who are 6 to 8 completed months, which results in small sample sizes and hence more imprecise estimates than for the larger sample.

**Table V.7. Impacts on Antenatal Care and Delivery Preparation, for Women Who Gave Birth in the Previous 9 Months (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| At Least 3 ANC Visits | 44.4 | 46.0 | 1.6 | 0.666 |
| At Least 2 TT Injections | 96.1 | 94.7 | -1.3 | 0.433 |
| IFA Tablets | | | | |
| At least 90 tablets received | 16.0 | 18.0 | 2.0 | 0.516 |
| At least 90 tablets consumed | 11.4 | 13.2 | 1.8 | 0.484 |
| Began taking tablets by month 4 | 31.3 | 36.4 | 5.0 | 0.181 |
| Transportation Plans | | | | |
| Obtained correct number of ambulance | 8.6 | 12.4 | 3.7 | 0.120 |
| Obtained number of private vehicle | 11.2 | 10.8 | -0.5 | 0.851 |
| Obtained number of FLW | 47.1 | 47.6 | 0.4 | 0.919 |
| Obtained any number | 55.4 | 54.5 | -0.8 | 0.829 |
| Delivery Preparations | | | | |
| Saved money for delivery | 86.3 | 86.7 | 0.4 | 0.858 |
| Identified facility for delivery or emergency | 55.8 | 61.7 | 5.9 | 0.205 |
| Identified person to accompany to facility | 55.8 | 57.5 | 1.7 | 0.707 |
| Discussed Delivery Plans with Husband | 93.1 | 92.2 | -0.9 | 0.652 |
| Discussed Delivery Plans with Mother-in-Law | 75.0 | 71.7 | -3.3 | 0.293 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters. Shaded outcomes are a focus of the TBGI intervention.

Sample size is 1,127 to 1,200.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

**Table V.8. Impacts on Delivery and Postpartum Care, for Women Who Gave Birth in the Previous 9 Months (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Facility Delivery | 78.6 | 82.4 | 3.8 | 0.182 |
| Cord Care | | | | |
| New blade | 94.0 | 92.5 | -1.5 | 0.472 |
| Clean thread | 93.9 | 94.4 | 0.5 | 0.739 |
| Nothing applied to cord | 56.0 | 55.6 | -0.4 | 0.928 |
| Clean-cord care[a] | 47.5 | 48.9 | 1.4 | 0.765 |
| Thermal Care | | | | |
| Bath delayed by at least 2 days | 46.8 | 47.8 | 0.9 | 0.802 |
| Skin-to-skin care | 50.1 | 51.6 | 1.5 | 0.669 |
| Early Initiation of Breastfeeding | | | | |
| Immediate breastfeeding | 56.4 | 58.4 | 2.0 | 0.613 |
| Child given prelacteals | 20.5 | 14.4 | -6.1 | 0.069* |
| Exclusive Breastfeeding | | | | |
| Exclusive breastfeeding for 6 months, children 6 months or older[b] | 26.3 | 34.2 | 8.0 | 0.145 |
| Exclusive breastfeeding in past 24 hours, children younger than 6 months[c] | 61.3 | 70.5 | 9.2 | 0.018** |
| Danger Signs | | | | |
| Experienced any danger sign during pregnancy or delivery | 60.2 | 52.4 | -7.8 | 0.114 |
| Among those who experienced danger signs: | | | | |
| Sought treatment | 68.6 | 64.7 | -3.9 | 0.334 |
| FLW advised to seek treatment | 31.2 | 33.7 | 2.5 | 0.605 |
| FLW advised where to go | 27.1 | 28.2 | 1.0 | 0.811 |
| Experienced any maternal danger sign in first 6 weeks | 40.1 | 37.6 | -2.5 | 0.591 |
| Among those who experienced danger signs: | | | | |
| Sought treatment for any sign | 57.8 | 62.0 | 4.2 | 0.518 |
| FLW advised to seek treatment | 21.8 | 31.3 | 9.5 | 0.056* |
| FLW advised where to go | 18.0 | 24.4 | 6.4 | 0.197 |
| Experienced any infant danger sign | 33.1 | 30.4 | -2.8 | 0.337 |
| Among those who experienced danger signs: | | | | |
| Sought treatment for any sign | 79.6 | 80.5 | 0.9 | 0.882 |
| FLW advised to seek treatment | 16.9 | 21.8 | 4.9 | 0.258 |
| FLW advised where to go | 12.4 | 15.6 | 3.2 | 0.324 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters. Shaded outcomes are a focus of the TBGI intervention.

Sample sizes are 1,046 to 1,200 (all women); 424 (children 6 months or older); 769 (children younger than 6 months); 692 (experienced danger signs during pregnancy or delivery); 504 (experienced danger signs in the first 6 weeks); and 396 (experienced infant danger signs).

[a]New blade and new thread used, and nothing applied to the cord.

[b]Based on self-reports of the duration of exclusive breastfeeding for children 6 months or older.

[c]Based on reports of liquids and solids fed to children younger than 6 months old in the previous 24 hours, following the recommended definition of the World Health Organization.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

Because the TBGI intervention focuses on the types of foods fed to children, we complemented our earlier analysis of child nutrition outcomes by examining impacts on food diversity in the previous 24 hours and food frequency in the previous seven days (Table V.9). To do so, we constructed indices of food diversity and food frequency based on reported consumption of different types of foods (slightly modified versions of the indices used by Garg and Chadha [2009]).

**Table V.9. Impacts on Child Nutrition, for Children 6 Months or Older (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Child Eats Solid or Semisolid Food | 62.4 | 67.4 | 5.0 | 0.211 |
| Child Began Eating Solid Food by Age 6 Months | 25.4 | 28.7 | 3.3 | 0.427 |
| Child Feeding in Previous Day | | | | |
| Times fed | 1.24 | 1.45 | 0.21 | 0.085* |
| Fed any meal from separate bowl | 31.7 | 40.9 | 9.2 | 0.047** |
| Fed any cereal-based meal | 46.3 | 55.3 | 9.0 | 0.016** |
| Amount fed (*katoris*) | 0.35 | 0.47 | 0.11 | 0.097* |
| Dietary Diversity Index (past 24 hours), Range 0–6[a] | | | | |
| Index = 0 | 49.2 | 40.7 | -8.5 | 0.009** |
| Index = 1 | 13.5 | 15.0 | 1.5 | 0.526 |
| Index = 2 | 25.1 | 27.3 | 2.3 | 0.538 |
| Index = 3–6 | 12.3 | 16.9 | 4.7 | 0.109 |
| | | | | 0.047*[u] |
| Mean | 1.05 | 1.25 | 0.20 | 0.032* |
| Food Frequency Index (past 7 days), Range 0–10[b] | | | | |
| Index = 0 | 41.8 | 37.2 | -4.5 | 0.251 |
| Index = 1–3 | 33.3 | 34.4 | 1.1 | 0.762 |
| Index = 4–10 | 24.9 | 28.4 | 3.5 | 0.384 |
| | | | | 0.505[u] |
| Mean | 2.05 | 2.26 | 0.21 | 0.259 |
| Anthropometry[c] | | | | |
| Stunted (< 2 standard deviations [s.d.]) | 26.5 | 27.0 | 0.5 | 0.897 |
| Wasted (< 2 s.d.) | 24.6 | 19.7 | -4.9 | 0.226 |
| Underweight (< 2 s.d.) | 29.5 | 25.9 | -3.6 | 0.333 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Tests for differences in distributions (age, birth parity, and SES quartile) were conducted using seemingly unrelated estimation regression in Stata. Sample includes 36 treatment and 36 control subcenters. Shaded outcomes are a focus of the TBGI intervention.

Sample size is 748 to 822.

[a]Index assigns one point for each of the following types of food eaten in the past 24 hours: rice, khichdi, or bread; daal; fruits and vegetables rich in vitamin A or dark green leafy vegetables; other fruits and vegetables; meat, fish, or eggs; and oil or ghee added to food (slightly modified from Garg and Chadha [2009]).

[b]Index assigns one point for each of the following types of food fed 1 to 3 times in the previous 7 days, and two points for each type of food fed 4 or more times in the previous 7 days: rice, khichdi, or bread; daal; fruits rich in vitamin A or dark green leafy vegetables; other fruits and vegetables; and meat, fish, or eggs (slightly modified from Garg and Chadha [2009]).

[c]Based on World Health Organization growth standards.

[d]*p*-values are for the test of equivalence of distributions.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

The findings suggest that there was no significant difference in terms of the frequency of feeding, but a marginally significant improvement in food diversity; the latter was driven by a decrease in the percentage of children who ate no nutritious foods in the previous day. Consistent with the modest impacts on complementary feeding practices, we also found little evidence of impacts on the standard anthropometric indicators of undernutrition, namely the proportion of children who were stunted (low height for age), wasted (low weight for height), or underweight (low weight for age). Stunting is a chronic condition that might have largely manifested by age 6 months and have been less responsive to complementary feeding, but one could potentially see lower rates of wasting and underweight children if there are large improvements in complementary feeding. Overall, the pattern of estimated impacts provides some evidence that the intervention had a positive effect on child nutrition, though only a handful of impacts are large and significant.

In addition to DPT3 immunizations, the focus of the TBGI goals, we examined the impact of the TBGI intervention on immunizations more generally (Table V.10). Again, we relied on information from child immunization cards and, when the card was blank or not available, from mothers' recall. Using the combined data from cards and self-reports, we found no statistically significant differences between the treatment and control groups in receipt of any of the specific routine immunizations, or in an overall measure of having received all routine immunizations (excluding measles). There was also no impact on a measure of timely vaccinations, namely whether the child received DPT3 by age 4 months. However, using information from immunization cards only, there was a significant impact on receipt of DPT1, which about 94 percent of the treatment group received versus 86 percent of the control group (a difference of about 8 percentage points, or 9 percent of the control mean). This could reflect the impact of a more general focus on encouraging immunizations by the FLWs. The card-only reports also suggested a larger impact of 6 percentage points in timely receipt of DPT3 (with 14 percent of the treatment group receiving DPT3 on time versus 8 percent in the control group), although this was only marginally significant. However, the card-only results should be viewed with some caution because treatment group members were more likely than their control counterparts to have a card (so that there might have been compositional differences between the treatment and control groups). Overall, the evidence that the TBGI intervention affected immunizations is therefore limited, although there is some suggestive evidence of improvements in some immunization outcomes.

Finally, although the TBGI goal in the reproductive health domain focused on the adoption of modern methods of contraception (either permanent or temporary methods) within six months of giving birth, we also explored impacts on contraceptive use for the broader sample. These results (Table V.11) suggest that the use of modern methods is significantly higher for women who gave birth *more* than six months ago. About 26 percent of these beneficiaries in the treatment group reported using these methods versus only 15 percent in the control group (a difference of almost 11 percentage points, or 69 percent of the control mean). Because use of permanent methods is not significantly different for these women, this is driven largely by an increase in the use of temporary modern methods. As a result of the strong positive impacts for these women, use of modern methods in the full sample is higher in the treatment group (about 19 percent versus 13 percent in the control group, a difference of 6 percentage points), though this impact is significant only at the 10 percent level. There was little evidence of impacts on other, more intermediate reproductive health outcomes that might affect contraceptive use, such as a woman's desire for another child or discussions on family planning with the woman's husband or mother-in-law in the previous three months.

**Table V.10. Impacts on Child Immunizations, for Children 6 Months or Older (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Child Ever Immunized | 98.0 | 98.9 | 0.9 | 0.306 |
| Vaccination Card Available, Among Children Ever Immunized | 51.3 | 60.2 | 9.0 | 0.045** |
| Card and Self-Reports | | | | |
| Received BCG | 94.5 | 94.4 | -0.1 | 0.964 |
| Received Polio 1 | 86.1 | 85.7 | -0.4 | 0.878 |
| Received Polio 2 | 77.4 | 79.6 | 2.2 | 0.479 |
| Received Polio 3 | 58.3 | 61.6 | 3.3 | 0.491 |
| Received DPT1 | 88.0 | 89.5 | 1.5 | 0.571 |
| Received DPT2 | 84.2 | 86.4 | 2.2 | 0.506 |
| Received DPT3 | 67.4 | 72.8 | 5.4 | 0.243 |
| Received DPT3 by age 4 months | 7.6 | 12.7 | 5.1 | 0.116 |
| Fully immunized (except measles) | 48.4 | 54.4 | 6.0 | 0.257 |
| Card Only | | | | |
| Received BCG | 92.6 | 94.6 | 2.0 | 0.389 |
| Received Polio 1 | 80.6 | 84.1 | 3.6 | 0.416 |
| Received Polio 2 | 71.8 | 77.8 | 6.0 | 0.183 |
| Received Polio 3 | 52.5 | 62.0 | 9.5 | 0.159 |
| Received DPT1 | 85.6 | 93.5 | 7.9 | 0.011** |
| Received DPT2 | 83.0 | 89.5 | 6.5 | 0.109 |
| Received DPT3 | 68.5 | 75.8 | 7.3 | 0.113 |
| Received DPT3 by age 4 months | 8.1 | 14.3 | 6.2 | 0.089* |
| Fully immunized (except measles) | 44.9 | 53.5 | 8.5 | 0.184 |
| Self-Reports Only | | | | |
| Received BCG | 91.6 | 89.2 | -2.4 | 0.484 |
| Received Polio 1 | 89.9 | 87.4 | -2.5 | 0.497 |
| Received Polio 2 | 81.7 | 81.6 | -0.2 | 0.968 |
| Received Polio 3 | 64.2 | 66.2 | 2.0 | 0.711 |
| Received DPT1 | 86.4 | 84.0 | -2.4 | 0.566 |
| Received DPT2 | 80.8 | 81.3 | 0.6 | 0.889 |
| Received DPT3 | 61.3 | 65.8 | 4.5 | 0.584 |
| Received DPT3 by age 4 months | 4.2 | -2.4 | -6.6 | 0.101 |
| Fully immunized (except measles) | 51.0 | 54.6 | 3.6 | 0.586 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters. Shaded outcomes are a focus of the TBGI intervention.

Sample sizes are 822 (all children), 801 (children ever immunized), 715 (card and self-reports), 468 (card only), and 261 (self reports only). Sample sizes for having received DPT3 by age 4 months are lower (542 for card and self-reports, and 88 for self-reports only).

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

**Table V.11. Impacts on Reproductive Health (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| Use of Contraception Among All Women | | | | |
| Use of permanent methods | 10.3 | 10.0 | -0.4 | 0.874 |
| Use of any modern method | 13.1 | 18.5 | 5.4 | 0.082* |
| Use of Contraception Among Women Who Gave Birth in Past 6 Months | | | | |
| Use of permanent methods | 7.6 | 4.6 | -3.0 | 0.073* |
| Use of any modern method | 10.5 | 10.3 | -0.2 | 0.957 |
| Use of Contraception Among Women Who Gave Birth More than 6 Months Ago | | | | |
| Use of permanent methods | 12.8 | 15.6 | 2.9 | 0.437 |
| Use of any modern method | 15.4 | 26.1 | 10.6 | 0.015** |
| Woman Would Like Another Child | 65.2 | 64.0 | -1.2 | 0.605 |
| Discussed Family Planning with Husband in Past 3 Months | 60.1 | 59.7 | -0.4 | 0.911 |
| Discussed Family Planning with Mother-in-Law in Past 3 Months | 37.6 | 37.4 | -0.2 | 0.964 |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment-control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters. Shaded outcomes are a focus of the TBGI intervention.

Sample sizes are 1,278 to 1,583 (all women); 767 (women who gave birth in past 6 months); and 816 (women who gave birth more than 6 months ago).

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

## G. Variation in Impacts, by Subgroup

To determine whether the overall impact of the TBGI intervention masked differential impacts for various demographic and socioeconomic subgroups, we estimated separate impacts for these subgroups by restricting the estimation sample accordingly. The key subgroups we examined were defined by scheduled caste or tribe status, socioeconomic quartile, literacy, and birth parity. For conciseness, we focused the subgroup analysis on seven outcomes most closely related to the TBGI goals (Table V.12). There was some suggestive evidence of a pattern of larger and more strongly significant positive impacts for women belonging to scheduled castes or tribes—especially for applying nothing to the cord, immediate breastfeeding, and feeding the child cereal-based complementary foods (although it is puzzling that some of the impacts on nonscheduled caste or tribe women are *negative* and statistically significant). However, in contrast, the impact on use of modern methods of contraception was larger and statistically significant for women belonging to nonscheduled castes or tribes. These results suggest that obtaining a better understanding of differences in the intervention's interactions with scheduled caste or tribe beneficiaries might be valuable if the intervention is scaled up (because these results were available only after we had completed the qualitative work, we did not explicitly explore this further in this study). For the other subgroups considered, there was no consistent pattern in the magnitude of the impacts. In addition, although impacts were statistically significant for certain outcome and subgroup combinations, the number of significant impacts for these other subgroups was similar to what one would expect by chance.

**Table V.12. Impacts for Beneficiaries with Greater Exposure to TBGI (percentages unless otherwise indicated)**

| | Control Mean | Adjusted Treatment Mean | Adjusted Difference | *p*-Value |
|---|---|---|---|---|
| **Pregnancy, Delivery, and Postpartum Care Outcomes: Women Who Gave Birth in Past 6 Months** | | | | |
| At Least 90 IFA Tablets Received | 15.8 | 14.1 | -1.7 | 0.596 |
| At Least 90 IFA Tablets Consumed | 10.2 | 10.7 | 0.5 | 0.864 |
| Transportation Plans | | | | |
| Obtained correct number of ambulance | 7.9 | 10.7 | 2.7 | 0.355 |
| Obtained number of private vehicle | 13.3 | 10.1 | -3.2 | 0.269 |
| Obtained number of FLW | 45.6 | 53.7 | 8.1 | 0.115 |
| Facility Delivery | 56.2 | 63.5 | 7.4 | 0.144 |
| Clean-Cord Care | 46.8 | 48.7 | 1.9 | 0.690 |
| Immediate Breastfeeding | 55.2 | 62.3 | 7.1 | 0.106 |
| FLW Gave Advice on IFA | 53.4 | 63.9 | 10.5 | 0.005** |
| FLW Visited in Last Trimester | 67.7 | 78.5 | 10.8 | 0.003** |
| FLW Visited in First 24 Hours After Delivery | 39.5 | 43.5 | 4.0 | 0.277 |
| **Other Outcomes: Women Who Gave Birth in Past 9 Months** | | | | |
| Child Eats Solid or Semisolid Food, Child 6 Months or Older | 46.2 | 47.6 | 1.4 | 0.818 |
| Child Feeding in Previous Day, Child 6 Months or Older | | | | |
| Times fed | 0.77 | 0.90 | 0.13 | 0.420 |
| Fed any meal from separate bowl | 20.3 | 27.1 | 6.8 | 0.185 |
| Fed any cereal-based meal | 30.9 | 35.1 | 4.2 | 0.459 |
| Amount fed (*katoris*) | 25.0 | 22.7 | -2.3 | 0.702 |
| FLW Visits Related to Complementary Feeding (child 5 months or older) | 24.1 | 33.4 | 9.3 | 0.055* |
| Child Received DPT3, Child 6 Months or Older | 50.3 | 56.2 | 5.9 | 0.186 |
| Uses Permanent Method of Contraception, All Women | 8.7 | 5.6 | -3.1 | 0.061* |
| Uses Modern Method of Contraception, All Women | 11.7 | 13.6 | 2.0 | 0.500 |
| FLW Visit to Discuss Family Planning | 18.5 | 26.2 | 7.7 | 0.011** |
| ASHA and AWW Ever Visited Together | 27.5 | 35.9 | 8.4 | 0.021** |
| ANM and ASHA/AWW Ever Visited Together | 22.5 | 29.9 | 7.4 | 0.036** |

Source: Ananya TBGI RCT Endline.

Notes: Treatment means and treatment–control differences are adjusted using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic); and subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters.

Sample sizes are 665 to 769 (top panel) and 398 to 532 (bottom panel). Sample sizes for visits by FLWs for family planning and joint visits in the bottom panel are higher (1,200).

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

**Table V.13. Impacts, by Subgroup**

| | Adjusted Difference (*p*-Value) | | | | | | |
|---|---|---|---|---|---|---|---|
| | Received 90 IFA Tablets | Any Transport Plan | Nothing Applied to Cord | Immediate Breast-feeding | Fed Any Cereal-Based Meal | Received DPT3, Child Older than 6 Months | Used Any Modern Method of Contraception |
| **Scheduled Caste/Tribe (SC/ST)** | | | | | | | |
| SC/ST | 0.9<br>(0.816) | 7.0<br>(0.232) | 11.5*<br>(0.092) | 10.3**<br>(0.009) | 11.0*<br>(0.056) | 6.8<br>(0.335) | 1.1<br>(0.767) |
| Non-SC/ST | 2.4<br>(0.419) | -7.5*<br>(0.079) | -8.6**<br>(0.031) | -3.0<br>(0.521) | 7.5<br>(0.143) | 4.9<br>(0.314) | 7.4**<br>(0.029) |
| **SES Quintile** | | | | | | | |
| Quintile 1 | -1.7<br>(0.724) | -0.4<br>(0.946) | 9.6<br>(0.230) | 18.6**<br>(0.007) | 6.7<br>(0.391) | -2.8<br>(0.730) | -1.7<br>(0.724) |
| Quintile 2 | 2.6<br>(0.588) | 6.5<br>(0.368) | -3.6<br>(0.575) | -2.5<br>(0.644) | 9.2<br>(0.297) | 7.0<br>(0.487) | 2.6<br>(0.588) |
| Quintile 3 | 1.7<br>(0.664) | -7.2<br>(0.190) | -0.7<br>(0.902) | 3.3<br>(0.521) | 12.8**<br>(0.040) | 10.5<br>(0.101) | 1.7<br>(0.664) |
| Quintile 4 | 2.1<br>(0.666) | -2.3<br>(0.703) | -5.0<br>(0.391) | -2.3<br>(0.683) | 9.1<br>(0.370) | 7.3<br>(0.186) | 2.1<br>(0.666) |
| **Literacy** | | | | | | | |
| Literate | 2.0<br>(0.646) | -0.8<br>(0.854) | 0.1<br>(0.981) | -1.1<br>(0.805) | 9.1*<br>(0.086) | -0.4<br>(0.929) | 2.0<br>(0.646) |
| Illiterate | 0.3<br>(0.903) | -2.6<br>(0.615) | -0.1<br>(0.977) | 2.9<br>(0.531) | 8.2<br>(0.137) | 10.9*<br>(0.091) | 0.3<br>(0.903) |
| **Birth Parity** | | | | | | | |
| 1 | 4.1<br>(0.300) | -0.7<br>(0.913) | 6.1<br>(0.381) | 1.9<br>(0.744) | 3.4<br>(0.652) | -7.5<br>(0.247) | 4.1<br>(0.300) |
| 2 | -0.2<br>(0.958) | 6.2<br>(0.315) | -5.0<br>(0.471) | 2.9<br>(0.579) | 18.0<br>(0.029) | 5.5<br>(0.468) | -0.2<br>(0.958) |
| 3 | 1.1<br>(0.846) | -7.5<br>(0.225) | 5.3<br>(0.519) | -9.7<br>(0.191) | 1.8<br>(0.846) | 18.4**<br>(0.042) | 1.1<br>(0.846) |
| 4 or more | -0.3<br>(0.951) | -7.5<br>(0.266) | 0.7<br>(0.913) | 8.0<br>(0.231) | 8.4<br>(0.356) | 7.0<br>(0.417) | -0.3<br>(0.951) |

Source: Ananya TBGI RCT Endline.

Notes: Adjusted differences are estimated using ordinary least squares regressions that control for stratum-fixed effects; indicators of demographic characteristics (SC/ST, Hindu, number of children, woman's age, woman's literacy, SES quartile, and indicators for missing values for each characteristic) other than those under consideration in a given panel of the table; and overall subcenter-level baseline means of the outcome (when available). Reported *p*-values account for clustering of standard errors at the subcenter level. Sample includes 36 treatment and 36 control subcenters.

Sample size is 150 to 1,051, and ranges are based on the outcome and subgroup considered.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

# VI. CONCLUSION AND IMPLICATIONS FOR SCALE-UP

Overall, the results suggest that after one year of implementation, the TBGI intervention improved the quantity and nature of FLW–household interactions relative to the standard Ananya package of interventions. Women who had recently given birth in the TBGI intervention areas were more likely than those in nonintervention areas to receive home visits from FLWs during their pregnancy and immediately after delivery, as well as visits specifically related to child feeding and reproductive health. Similarly, they were more likely than those in nonintervention areas to have received advice on key maternal and child health topics (including but not limited to topics related to the TBGI goals), to have been exposed to Ananya behavior change tools, and to have received joint visits from FLWs. The results from the FLW surveys suggest that these changes were facilitated by a substantial change in coordination among FLWs in treatment subcenters, in keeping with the team-based focus of the intervention. The TBGI intervention therefore appears to have led to an increased effort from FLWs to change beneficiaries' behaviors to achieve their team goals, with positive spillover effects on their interactions with households more generally.

However, the impact of the TBGI intervention on women's health-related knowledge and behaviors within the one-year evaluation period was more modest, even for outcomes directly related to the TBGI goals. Although the study was powered to detect changes in the 5- to 10-percentage point range for the full sample, most differences in these outcomes were smaller. Nevertheless, there was some evidence of positive effects for certain outcomes, particularly for exclusive breastfeeding, age-appropriate complementary feeding, use of modern contraceptive methods, and certain child immunization outcomes. These results suggest that TBGI has the potential to add value to the existing Ananya interventions and lead to larger changes in household behaviors at relatively low additional cost (given the low cost of the gifts provided as incentives), though translating improved FLW–household interactions from TBGI into major behavior change is an ongoing challenge.

We cannot determine with certainty whether we would find similar impacts of the TBGI intervention in the longer term in its current form, or if it is scaled up more broadly. For example, it is possible that FLW teamwork would become more efficient and effective over time, leading to larger impacts on beneficiaries' behaviors. Alternatively, the initial enthusiasm could wear off, so that the impacts that we found would not be sustained. Similarly, our results do not necessarily generalize to the intervention if it were scaled up throughout Bihar as part of a scaled-up Ananya program, because impacts could differ in contexts outside of Begusarai district and because implementation could change if the government, rather than CARE, implemented the intervention (for example, CARE staff would not be available to conduct subcenter meetings).

Nevertheless, the findings from our process study and the FLW reports provide some inputs for sustainability if the program is scaled up. First, the process study findings, corroborated by the ANMs' own reports, suggest that for this intervention to be led by ANMs, it will be important to further train and strengthen their role so they can lead the subcenter meetings, including communication of content to more technical information of how to calculate targets. This is critical to scale-up, because maintaining the high level of involvement by CARE staff is unlikely to be sustainable on a larger scale. Second, procurement mechanisms within the government could alter how the incentive structure for the FLWs is implemented relative to this pilot, in which CARE handled the procurement of the relatively small number of gifts required. The procurement and logistical challenges to implement the incentive scheme at scale could pose challenges to the government to duplicate this process. Third, the FLWs might benefit with more guidance on defining what it means to achieve goals and devising a stringent monitoring mechanism to track this

achievement. Both our qualitative and survey data show that FLWs articulated a varied understanding of what accomplishing each of the seven goals meant. Uncertainty around these definitions might undermine the FLWs' ability to effect changes in beneficiaries' behaviors regarding key maternal and child health practices. Further, although the program seemed to lead to more home visits and improvement in some outcomes with little external monitoring, it is possible that, over time, the model can incentivize FLWs to over-report their achievements. However, any monitoring mechanism identified will also have to be easily managed by the government for sustainability of the intervention.

# REFERENCES

Basinga, Paulin, Paul Gertler, Agnes Binagwaho, Agnes L.B. Soucat, Jennifer R. Sturdy, and Christel Vermeersch. "Paying Primary Health Care Centers for Performance in Rwanda." World Bank Policy Research Working Paper Series, no. 5190. Washington, DC: World Bank, 2010.

Dieleman, Marjolein, and Jan Willem Harnmeijer. "Improving Health Worker Performance: in Search of Promising Practices". Geneva, Switzerland: World Health Organization, September 2006.

Eichler, Rena, Ruth Levine, and the Center for Global Development PBI Working Group. *Performance Incentives for Global Health: Potential and Pitfalls*. Washington, DC: Brookings Institution Press, 2009.

Garg, Aashima, and Ravinder Chadha. "Index for Measuring the Quality of Complementary Feeding Practices in Rural India." *Journal of Health, Population, and Nutrition,* vol. 27, no. 6, 2009, p.763.

Herzberg, Frederick. "One More Time: How Do You Motivate Employees?" *Harvard Business Review,* September–October 1987, pp. 5–16.

Lim, Stephen S., Lalit Dandona, Joseph A. Hoisington, Spencer L. James, Margaret C. Hogan, and Emmanuela Gakidou. "India's Janani Suraksha Yojana, a Conditional Cash Transfer Programme to Increase Births in Health Facilities: An Impact Evaluation." *The Lancet,* vol. 375, no. 9730, 2010, pp. 2009–2023.

Mazumdar, S., A. Mills, and T. Powell-Jackson. "Financial Incentives in Health: New Evidence from India's Janani Suraksha Yojana." Working paper. London: Oxford University, 2011.

Miller, Grant, and Kimberly Singer Babiarz. "Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs." NBER Working Paper No. 18932. Cambridge, MA: National Bureau of Economic Research, 2013.

Peabody, John W., Riti Shimkhada, Stella Quimbo, Orville Solon, Xylee Javier, and Charles McCulloch. "The Impact of Performance Incentives on Child Health Outcomes: Results from a Cluster Randomized Controlled Trial in the Philippines." *Health Policy and Planning,* vol. 28, no. 8, 2013.

Singh, Prakarsh. "Performance Pay and Information: Reducing Child Malnutrition in Urban Slums." Working paper. Amherst, MA: Amherst College, 2011.

Wang, Hong, Rajini K. Juyal, Sara A. Miner, and Elizabeth Fischer. "Performance Based Performance System for ASHAs in India: What Does International Experience Tell Us?" Bethesda, MD: The Vistaar Project, IntraHealth International, Abt Associates, March 2012.

World Health Organization (WHO). "Indicators for Assessing Infant and Young Child Feeding Practices. Part II: Measurement." WHO Press: Geneva, Switzerland, 2010. Available at http://whqlibdoc.who.int/publications/2010/9789241599290_eng.pdf.

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

# APPENDIX A

# STUDY DESIGN, SAMPLING, AND ANALYTIC APPROACH

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

## A. Study Design

As described in the text, we conducted a clustered randomized controlled trial (RCT) to evaluate the impacts of the Team-Based Goals and Performance-Based Incentives (TBGI) intervention. We conducted randomization at the subcenter level, using all the subcenters in five selected blocks in Begusarai district. Specifically, we randomized the 76 subcenters in these blocks into equal-sized treatment and control groups using a stratified random assignment procedure that involved organizing the subcenters in each block into strata based on the number of Anganwadi Centers (AWCs) served by the subcenter (a proxy for the size of the population served). Specifically, we divided subcenters in each block into a stratum of "small" subcenters and a stratum of "large" subcenters.[20] We then conducted the randomization separately in each stratum by assigning half the subcenters in the stratum to treatment and half to control (or about half in the case of strata with an odd number of subcenters) (Table A.1). The stratification procedure ensured that the treatment and control groups were balanced by the size of population served and reduced variance (hence improving statistical power) in the analysis.[21]

**Table A.1. Stratified Randomization of Subcenters in Begusarai**

| Stratum Number | Block | Number of AWCs Served | Total Number of Subcenters | Treatment Subcenters | Control Subcenters |
|---|---|---|---|---|---|
| 1 | Bachwara | < 6 | 11 | 5 | 6 |
| 2 | Bachwara | ≥ 6 | 9 | 5 | 4 |
| 3 | Chaurahi | < 6 | 19 | 9 | 10 |
| 4 | Bhagwanpur | < 6 | 6 | 3 | 3 |
| 5 | Bhagwanpur | ≥ 6 | 11 | 6 | 5 |
| 6 | Birpur | ≥ 6 | 8 | 4 | 4 |
| 7 | Nawkothi | < 6 | 6 | 3 | 3 |
| 8 | Nawkothi | ≥ 6 | 6 | 3 | 3 |
| **Total** | | | **76** | **38** | **38** |

Note: The table shows the allocation of subcenters in selected blocks in Begusarai to treatment and control groups. All subcenters in the selected blocks were included. Subcenters were organized into strata, which are defined by block and the number of AWCs served (above or below the median of 6). Random assignment was conducted separately within each stratum to obtain an approximately equal allocation to the treatment and control groups.

## B. Sampling Approach

### 1. Sampling for Beneficiary Data Collection

To obtain a sample frame of eligible beneficiaries in the communities served by the treatment and control subcenters (the target population), we had to conduct a listing exercise that recorded information on all birth events in the previous 12 months. To keep this listing exercise manageable, we randomly selected two villages served by each subcenter for inclusion in our surveys, using a list

[20] Because the median subcenter in our sample served six AWCs, we used this as the cutoff to distinguish between small and large subcenters. Some blocks consisted entirely of small or large subcenters using this cutoff; in these cases, the entire block was used as a single stratum.

[21] Given the limited information available on subcenter characteristics (block and size) before randomization, we deemed this crude stratification scheme to be sufficient; it seemed unlikely that a more elaborate stratification scheme using this information (for example, forming smaller size strata within each block) would be advantageous.

of villages linked to each subcenter provided by CARE's field team. If a large village (one with a population of 150 or more, as identified by CARE) was selected, we then organized the village into approximately equal-sized segments (75 to 150 households per segment) and randomly selected one segment of the village for surveying.[22]

We first intended to survey all eligible beneficiaries in the selected villages (or segments); based on existing data about birth rates, we expected our sampling approach to meet our sample size target of about 20 eligible beneficiaries per subcenter. However, after conducting the listing, we discovered that some segments had many more eligible beneficiaries than we expected, which would have increased the time and cost of data collection without much benefit in terms of statistical power. Therefore, if the listing identified 30 or more eligible beneficiaries for a subcenter (across both of the selected villages or segments), we randomly selected 25 of these into the sample (by randomly selecting an approximately equal number in each village or segment).

Overall, the beneficiary sampling approach in each subcenter consisted of up to three stages, involving the sequential selection of villages, segments, and beneficiaries. At endline, we returned to the same villages (or segments) but conducted a new listing to identify women who had given birth in the previous 12 months (again, we drew samples when necessary to meet our targeted sample sizes per subcenter).

### 2. Sampling for Frontline Worker Data Collection

Our overall goal was to draw a representative sample of frontline workers (FLWs)—including accredited social health activists (ASHAs), Anganwadi workers (AWWs), and auxiliary nurse midwives (ANMs)—linked to the subcenters in our sample. The targeted sample size based on our statistical power calculations was nine FLWs per subcenter, including four ASHAs, four AWWs, and one ANM. For the baseline ASHA and AWW samples, we used a village-wise list (provided by CARE's field team) of ASHAs and AWWs in each subcenter. Rather than simply draw a random sample of ASHAs and AWWs from this list, we decided to focus our sample of ASHAs and AWWs on the villages selected for the household survey. This was both to gain logistical simplicity and to have a closer link between household and FLW outcomes.

We therefore attempted to meet our targeted sample size of four ASHAs and four AWWs per subcenter by drawing a random sample of ASHAs and AWWs in the villages selected for the household sample in each subcenter. However, because some villages did not have sufficient ASHAs or AWWs available (for example, due to unfulfilled vacancies), it was sometimes necessary to select additional ASHAs or AWWs from other villages to reach the targeted sample size for the subcenter. To follow the same sampling approach in these cases, we first randomly selected an additional village(s) in the relevant subcenter, and then randomly selected the appropriate number of ASHAs or AWWs from this village or these villages.

Although this provided us with a representative sample of ASHAs and AWWs, we also wanted to ensure that we surveyed the ASHAs and AWWs linked to the particular women in our beneficiary sample. Because some villages were large and we selected only one segment for the household survey, the beneficiaries that we randomly sampled in the selected segment might not have been

[22] When the subcenter served only a single (usually large) village, we selected two segments from this village.

served by the ASHA or AWW (there was no way for us to know *ex ante* which FLWs served which segments). We therefore asked surveyed beneficiaries to provide us with the names of the ASHAs and AWWs serving them, and added these to the FLW sample if they were not already included.

The final baseline ASHA and AWW samples therefore included the following groups: (1) FLWs randomly drawn from the same villages selected for the household sample; (2) FLWs randomly drawn from other villages linked to the randomized subcenters; and (3) FLWs added *ex post* to the sample because they served the women sampled for the beneficiary survey, but had not been selected into the FLW sample. Finally, we used a randomly ordered replacement list (prioritizing ASHAs and AWWs in villages already selected) to ensure that we met our sample size targets, despite the fact that some selected ASHAs or AWWs could not be found or refused to participate. At endline, we attempted to survey all the ASHAs and AWWs identified at baseline, plus any new ASHAs and AWWs serving the sampled beneficiaries (there was relatively little turnover—we identified only 4 new ASHAs and 5 new AWWs out of a sample of about 300 of each).

As noted earlier, we attempted to survey all ANMs in the treatment and control subcenters at baseline and endline, respectively (most subcenters had one ANM, but some had two). There was substantial turnover in the ANMs between baseline and endline, so that the two samples were quite different (30 of the 109 ANMs identified at baseline were no longer in their position at endline).

## C. Analytic Approach

### 1. Beneficiary Surveys

Because randomization should ensure that the treatment and control groups are similar in all respects other than receipt of the intervention, we could have estimated impacts simply by computing the difference in mean outcomes between the two groups at endline. However, we instead estimated impacts in a regression framework, using the following regression model for beneficiaries in our sample:[23, 24]

$$(1) \quad Y_{ijk,post} = \alpha + \beta T_{jk} + \alpha_k \lambda_k + \gamma X_{ijk} + \delta Z_{jk} + \mu_{jk} + \varepsilon_{ijk}$$

where $Y_{ijk,post}$ is the outcome for woman $i$ in subcenter $j$ in stratum $k$ at endline; $T_{jk}$ is a binary indicator for subcenter $j$ being in the treatment group; $\lambda_k$ is a vector of stratum indicators (one for each random assignment stratum); $X_{ijk}$ and $Z_{jk}$ are vectors of individual- and subcenter-level covariates, respectively, that could be related to the outcome of interest (individual-level covariates include household demographics such as whether a woman belongs to a scheduled caste or tribe, is Muslim, the number of children, age, literacy, and socioeconomic status [SES] quartile, whereas the

[23] Using a regression framework enabled us to explicitly account for the method of randomization by induding stratum-level indicators. By controlling for additional individual and baseline subcenter-level characteristics, we were also able to reduce the variance in the outcome (and hence increase statistical power) and control for differences between the treatment and control groups that could have arisen by chance. As a robustness check, we compared our results with those from simple unadjusted treatment–control comparisons; the results were largely similar.

[24] In the case of binary outcomes such as training receipt or employment, Equation (1) is termed a linear probability model. Although probit or logit models are often used for binary outcomes, we prefer the linear probability model because it is easier to interpret and relies on weaker parametric assumptions. In practice, the probit or logit and linear probability models yield similar results for the estimated impacts.

subcenter-level covariates include the subcenter-level mean of the outcome, when available); $\mu_{jk}$ is a subcenter-level error term; and $\varepsilon_{ijk}$ is an individual error term. The coefficient of interest is $\beta$, which gives the impact of the intervention on the outcome of interest; conceptually, this is only the regression-adjusted difference in mean outcomes between the treatment and control groups. All woman-level regressions were weighted to account for differing sampling probabilities and to ensure treatment–control balance within random assignment strata; the estimated impacts can therefore be interpreted as the impacts for the average woman in the treatment subcenters. We adjusted all standard errors to account for the correlation in outcomes among beneficiaries linked to the same subcenter using the cluster adjustment in Stata.

## 2. FLW Surveys

We used a similar regression framework to determine the impact of the interventions on outcomes for ASHAs and AWWs.[25] In all analyses, we pooled the data collected from workers in these two cadres to estimate the following regression model:[26]

$$(2) \quad \begin{aligned} Y_{ijk,post} &= \alpha + \beta T_{jk} + ASHA_{ijk} + \alpha_{1k}\lambda_k \\ &+ \alpha_{2k}\lambda_k * ASHA_{ijk} + \gamma_1 X_{ijk} + \gamma_2 X_{ijk} * ASHA_{ijk} + \delta Z_{jk} + \mu_{jk} + \varepsilon_{ijk} \end{aligned}$$

Similar to the household analysis, $Y_{ijk,post}$ is the outcome for FLW $i$ in subcenter $j$ in stratum $k$ at endline; $T_{jk}$ is a binary indicator for subcenter $j$ being in the treatment group; $ASHA_{ijk}$ is an indicator for the FLW being an ASHA (rather than an AWW); $\lambda_k$ is a vector of stratum indicators (one for each random assignment stratum); $X_{ijk}$ and $Z_{jk}$ are vectors of FLW- and subcenter-level covariates, respectively, that could be related to the outcome of interest; $\mu_{jk}$ is a subcenter-level error term; and $\varepsilon_{ijk}$ is an individual error term. Note that we allow stratum indicators and FLW characteristics to differentially affect outcomes for ASHAs and AWWs as sampling occurred separately by cadre and certain characteristics might be differentially helpful to ASHAs or AWWs in trying to improve health-related outcomes. However, we restrict the regression so that subcenter characteristics have the same impact on outcomes measured in the ASHA and AWW surveys; these characteristics are included in the regression to control for a community's influence on the outcome, which should not vary by FLW cadre. Subcenter characteristics are taken from beneficiaries in the endline household survey and include the share of beneficiaries who are literate, Muslim, in a scheduled caste or scheduled tribe, and first-time mothers. The FLW characteristics include controls for worker's age, scheduled caste or scheduled tribe status, religion, education level, and an indicator for whether the FLW serves a village less than one hour travel time from her assigned subcenter.

[25] All regressions using data from the ASHA and AWW surveys were weighted, correcting for differing sampling probabilities and ensuring treatment–control balance. Weights were normalized so that the estimated effects represent the average FLW. In addition, in all FLW regressions, we adjust for clustering as in the household-level analysis, using the cluster adjustment in Stata at the subcenter level.

[26] We also analyzed several key outcomes for ASHAs and AWWs separately but found few results that varied across cadres.

We estimated a more parsimonious model to examine the impacts of TBGI on ANMs:

(3) $$Y_{ijk,post} = \alpha + \beta T_{jk} + \alpha_k \lambda_k + \mu_{jk} + \varepsilon_{ijk},$$

with variables defined as in the other FLW regressions.[27] Because our analysis was not designed to have the statistical power to distinguish impacts within the ANM sample (when not pooled with other FLW cadres), these results should be thought of as providing descriptive context for the ASHA, AWW, and household results. Because of this descriptive nature, the model excludes any subcenter or ANM characteristics.

[27] Because almost all ANMs in treatment and control subcenters are included in our analysis, we did not weight the ANM data. As in the ASHA/AWW regressions, we adjust for clustering using the cluster adjustment in Stata at the subcenter level.

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

# APPENDIX B

# DETERMINING WHEN TARGETS ARE MET

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

**Table B.1. Determining When Targets Are Met**

| | Mean |
|---|---|
| When is transportation goal met? | |
| After informing household | 0.420 |
| After giving FLW number | 0.418 |
| After giving ambulance number | 0.489 |
| After someone at household informed you | 0.303 |
| After observing information in household | 0.223 |
| After accompanying family to PHC for delivery | 0.365 |
| Other FLW is responsible for this target | 0.052 |
| When is 90+ IFA tablet goal met? | |
| After giving tablets | 0.354 |
| After asking if received tablets | 0.270 |
| After asking if consumed tablets | 0.700 |
| After asking to see empty strips | 0.579 |
| After asking about stool color | 0.429 |
| Other FLW is responsible for this target | 0.033 |
| After seeing her tongue's color | 0.043 |
| When is immediate breastfeeding goal met? | |
| After informing mother | 0.321 |
| After asking mother | 0.577 |
| After observing if she breastfed | 0.716 |
| After asking other FLW | 0.274 |
| Other FLW is responsible for this target | 0.027 |
| When is clean cord goal met? | |
| After informing mother | 0.404 |
| After attending birth and observing | 0.520 |
| After asking mother | 0.549 |
| After observing clean cord | 0.426 |
| Other FLW is responsible for this target | 0.040 |
| When is complimentary feeding goal met? | |
| After informing mother | 0.490 |
| After asking mother about types and quantities of food | 0.522 |
| After observing child eating correct food | 0.667 |
| After checking child's health at home visits | 0.358 |
| Other FLW is responsible for this target | 0.033 |
| When is family planning goal met? | |
| After informing mother about methods | 0.449 |
| After distributing condoms and pills | 0.577 |
| After asking about TL and IUDs | 0.543 |
| After asking about use | 0.577 |
| Other FLW is responsible for this target | 0.031 |
| When is DPT-3 goal met? | |
| After informing mother about importance | 0.347 |
| After telling them to complete all immunizations | 0.299 |
| After asking them to come to immunization day | 0.349 |
| After asking if child got DPT-3 | 0.540 |
| After checking ANM's register | 0.289 |
| After observing child get shot | 0.491 |
| Other FLW or ANM is responsible for this target | 0.019 |

Source: Ananya TBGI RCT Endline, treated sub-centers only. All means are weighted.

Note: The joint sample size of ASHAs and AWWs is 273. All means are weighted.

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

Mathematica® is a registered trademark of Mathematica Policy Research