# A mixed-methods evaluation to determine the effects of a novel mHealth platform for maternal child health tracking in rural Udaipur, India

Ruchit Nagar
Harvard Medical School, Khushi Baby

Mohammad Sarparajul Ambiya
Khushi Baby

Pawan Singh
Khushi Baby

Hamid Abdullah
Khushi Baby

Vijendra Banshiwal
Khushi Baby

Logan Stone
University of Miami Miller School of Medicine, Khushi Baby

Deepa Manjanatha
Harvard School of Public Health, Khushi Baby

Preethi Venkat
Georgetown Medical School, Khushi Baby

Dhanwant Purawat
Khushi Baby

Vaidehi Supatkar
Khushi Baby

Arjun Singh
Khushi Baby

Saachi Dalal
Khushi Baby

Mohammed Shahnawaz
Khushi Baby

**Grantee Final Report**

**Accepted by 3ie: June 2020**

International Initiative for Impact Evaluation

## Note to readers

This impact evaluation has been submitted in fulfilment of requirements under grant TW10.1078 issued under Innovations in Increasing Immunisation Evidence Programme. This version is being published online as it was received. A copy-edited and formatted version will be available in the 3ie Impact Evaluation Report Series in the near future.



The 3ie technical quality assurance team comprises Monica Jain, Avantika Bagai, Kirthi V Rao, Sayak Khatua, Manya Malik, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder.

Funding for this impact evaluation was provided by Bill & Melinda Gates Foundation. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: Nagar, R, Mohammad, SA, Singh, P, Abdullah, H, Banshiwal V, Stone, L, Manjanatha, D, Venkat, P, Purawat D, Supatkar, V, Singh, A and Shahnawaz, M, 2020. *A mixed-methods evaluation to determine the effects of a novel mHealth platform for maternal child health tracking in rural Udaipur, India,* 3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation (3ie).

.

## Acknowledgements

We would like to thank the Udaipur District Health Society – the 87 ANMs and 310 ASHAs who worked closely with us, the RCHO officers, and especially the now Deputy Director of Health Udaipur, Dr. Sanjeev Tak for making this project a reality. We would like to thank 3ie for providing us the funds to launch not only our project for this impact evaluation but in effect, our organization, by providing financial runway for over two years. Specific thanks should be given to Dr. Monica Jain for her valuable inputs and to our colleagues in the TW10 series for sharing their insights.

We would like to thank our other funders who have made our work possible: GAVI, Johnson and Johnson, UNICEF UK, ARM, ISIF Asia, InnovateHealth Yale, John Hopkins Future Health Systems, Kickstarter backers, Bitgiving backers, Aquent Design for Good, the Human Nature Lab at Yale, and Davenport College at Yale.

We would like to thank our partners who have supported us in-kind: UNICEF Innovation, Cooley (legal), Idemia (hardware), and Mobisoft Infotech (software development). Without your support, we would not be up and operational today.

We would like to thank our advocates at GAVI and Johnson and Johnson for giving us an opportunity to dream bigger about where KB might be headed.

We would like to thank our advisors and mentors including Bo Hopkins, Dr. Joe Zinter, Dr. Zeena Johar, Shail Sinhasane, Blair Palmer and Dr. Bhupendra Tripathi, for their steady support, our global family of team members for their dedication, and our families for putting up with us during this long journey (and the journey ahead).

We would like to thank our technology team – Rahul Kumar Das, Arul Srinivasan, Sanjay Kumar, Tanooj Luthra, Jitendra Kolate, Rishi Rawat, Sukumar Salokhe, and Gurudeep Singh for the challenge that they have taken upon to rebuild and maintain the KB platform.

We would like to thank our interns and volunteers for contributing their insight and skills into our process, with special acknowledgements to Sara Locke, Farhana Jamal, and Kyle Engel.

Most importantly, we would like to thank our team of Core Staff and KB Monitors on the ground who have worked tirelessly and sacrificed incredibly over the past year to make this project and organization a reality. Specifically we would ilke to thank Navratan Pargi, Prem Shankar Meena, Prabhu lal Gorna, Sukh Lal Dungari, Ravi Singh, Md. Nabi Hussain, Mukesh Trivedi Neha Purawat, Danny Kumar Purbia, Bhupendra Singh Chouhan, Bhavishya Purbia, and Ashok Kumar Nama.

# Contents

## List of figures and tables

## Abbreviations/Acronyms

| | |
|---|---|
| ANC | Antenatal Care Checkup / Antenatal Care |
| ANM | Auxiliary Nurse Midwife |
| ANMOL | ANM Online |
| ASHA | Accredited Social Health Activist |
| BCMO | Block Chief Medical Officer |
| BCG | bacille Calmette-Guerin (a vaccine for Tuberculosis) |
| BPM | Block Program Manager |
| CHC | Community Health Centre |
| CMHO | Chief Medical and Health Officer (District) |
| DEO | Data Entry Operator |
| 3IE | International Initiative for Impact Evaluation |
| EBF | Exclusive Breast Feeding |
| EDD | Expected Date of Delivery |
| HER | Electronic Health Record |
| FIC | Full Immunization Coverage |
| FICT | Full Immunization Coverage and Timeliness |
| GAVI | Global Alliance for Vaccines Initiative |
| GOI | Government of India |
| IFA | Iron Folic Acid |
| KB | Khushi Baby |
| KPI | Key Performance indicator |
| LHV | Lady Health Visitor |
| MCTS | Maternal and Child Tracking System |
| MOIC | Medical Officer in Charge |
| MCH | Maternal Child Health |
| MNCH | Maternal Neonatal and Child Health |
| MOHFW | Ministry of Health and Family Welfare |

| | |
|---|---|
| MUAC | Mid-upper-arm Circumference |
| NFC | Near Field Communication |
| NFHS | National Family Health Survey |
| NGO | Non-Governmental Organization |
| NHM | National Health Mission |
| OPV | Oral Polio Vaccine |
| OBC | Other Backward Caste |
| PENTA | Pentavalent Vaccine |
| PCTS | Pregnant Women and Child Health Tracking System |
| PHC | Primary Health Centre |
| PNC | Postnatal Care |
| RCH | Reproductive and Child Health |
| SC | Sub Centre |
| SC | Scheduled Caste |
| ST | Scheduled Tribe |
| TT | Tetanus Toxoid Vaccine |
| UI/UX | User interface / user experience |
| VHND | Village Health and Nutrition Day |

## Executive summary

Over 500,000 children die every year in India from vaccine preventable disease. Khushi Baby is a solution built to change the way maternal and child health care is tracked and engaged at the last mile. Our goal is to understand which mothers and children are falling through the cracks.

The Khushi Baby (KB) system centers around a novel health record for patients – the Khushi Baby Pendant. This health record is digital, wearable, culturally symobolic, durable, biometrically-secured, and costs less than $1. Health providers use the Khushi Baby mobile application to scan the Khushi Baby Pendant, which allows them to both read, and update the medical history, at the last mile, *without the need of prior data synchronization*. The Khushi Baby App also acts as a decision-support tool for frontline health workers. Data from the application is synced to the Khushi Baby Dashboard for health officials to see high risk reports, supply side gaps, and health worker targets. The dashboard sends automated voice call reminders to individual families in the local dialect. Health workers and health officials discuss repors on Khushi Baby WhatsApp groups and Khushi Baby Monitors provide in-field support to health workers for follow-up visits.

The International Initiative for Impact Evaluation (3ie) awarded Khushi Baby a grant to evaluate the impact of the platform, which has been used to track the health of over 25,000 mothers and infants to date. Through a rigorous, randomized controlled trial over the past two years, the Khushi Baby team has followed over 3000 mothers and conducted nearly 10,000 interviews of mothers, frontline health workers, health supervisors, and health officials. The unadjusted results of the trial showed Khushi Baby increased full infant immunization rates by a factor of 2.03 and decreased the infant moderate acute malnutrition rates by a factor of 0.23. Put differently, Khushi Baby increased full immunization rates by 12.2 percentage points and decrased infant acute malnutrition rates by 4 percentage points. ***After adjustment for confounders, the final results showed the Khushi Baby system increased full infant immunization by a factor of 1.66 and decreased acute malnutrition rates by a factor of 0.26, both statistically significant and politically meaningful increases, when compared to the existing paper-based tracking system. The cost per 10 percentage point increase in likelihood of full immunization was $0.68 (50 INR) per beneficiary.***

The KB system retrieved data from the field in 4 hours, at an average cost of $4.47 or just under 300 INR per additional beneficiary per year (inclusive of all costs of technology and human resources), and with a 4/5 star approval from the frontline health workers. The results were not statistically significant for decreasing hospitalization rates or infant mortality.

These findings are particularly noteworthy for three reasons. First, the beneficiaries (mothers) receiving this intervention were only exposed after delivering their child. However, the KB App is also used for tracking of antenatal care with automated community engagement of pregnant women as well. This additional early tracking and intervention may have resulted in even higher health behaviour outcomes during childhood, not measured in this trial. Second, these results were found despite frontline health workers still having to conduct double-work on paper registers to comply with mandated reporting protocols established by the State of Rajasthan. Third, these results are impressive considering the technical, financial, political and operational challenges faced to establish the organization and build the intervention and deploy the intervention at the last mile. As a result, and with the support of Dr. Sanjeev Tak, ex-Chief Medical & Health Officer of the Udaipur District, the Khushi Baby intervention was given clearance to scale from 350 villages in less than 5 administrative blocks to cover over 1000 villages across the 5 blocks universally and to cover the entire District's maternal and child health (MCH) tracking and engagement operation by 2020. This scale-up is underway, and additionally the State Department of Health and Family Welfare has selected Khushi Baby's platform as a base to scale-up state-wide in multiple parallel districts in the coming year.

## Introduction

According to the Ministry of Health and Family Welfare, India sees an estimated 500,000 children die annually of vaccine preventable disease.[1] Despite much progress over the past decade, still nearly 60,000 mothers die annually in India from preventable causes related to pregnancy and childbirth.[2] A variety of cost effective solutions exist to avert this unnecessary maternal, neonatal, and child mortality: immunizations, treatment of febrile illness, simple practices for newborn care, and micronutrient supplementation for pregnant women stand out as successful interventions along the continuum of *ante*-, *intra*-, and *post*- natal care. Increasing coverage of these essential services requires further attention.[3]

High IMR and MMR in India is additionally concerning due to unreliable estimates from health surveillance systems. In the state of Rajasthan, 2.3 Million mothers registered during pregnancy were lost to follow-up in the state e-health registry between 2011-2016 according to a report from the Comptroller and Auditor General of India.[4] Without first knowing which children (and mothers) are being left out, strategies to drive behavior change for better uptake of essential MCH services would likely be nonspecific and ineffective.

Consider the case of Udaipur, a “medium-focus” district for immunization coverage improvements in southwest Rajasthan.[5] The Annual Health Survey 2012-13 reported full immunization coverage rates of 79.8% in rural areas for those children 12-23 months of age.[6] On the contrary, the reputable National Family Household Survey conducted in 2014 found just 37.2% of children in rural areas fully immunized in the same district.[7] For the entire district of Udaipur, Rajasthan’s Pregnant Woman and Child Tracking System (PCTS) showed 44.5% of children age 12-23 months fully immunized in 2013.[8] Differences in denominators, sampling strategies, and data quality might have resulted in this wide range of estimates. Multiple handoffs within the PCTS tracking system may have also contributed to discrepancies [1].

Deficits in data quality noted at the State and District level reflect the current state of tracking at the grassroots level - the villages, where failure to account for mothers and children results in the aforementioned estimates of IMR and MMR. The purpose of this evaluation is to rigorously investigate the merit of a novel, culturally-tailored, data-vigilant, mHealth platform for rural MCH tracking in India: Khushi Baby. Specifically, this evaluation addresses the effectiveness of a systematic multi-stakeholder, multi-component intervention on: improvement in MCH data

---

[1]Mission Indradhanush <http://www.missionindradhanush.in/about.html>
[2]Unicef India. Maternal Health. <http://unicef.in/Whatwedo/1/Maternal-Health>
[3]Horton, S and Levin, C. Chapter 17. <https://www.ncbi.nlm.nih.gov/books/NBK361909/>
[4]Goswami R. 2.3 million pregnant women in Rajasthan ‘missing’ between 2011 and 2016. Hindustan Times. May 1, 2017.
[5] Medium focus district label was provided under the Mission Indradhanush campaign Phase II, a national campaign to improve child immunization coverage.
[6]Annual Health Survey 2012-13
[7]National Family Household Survey 2014
[8]PCTS

reliability, MCH data retention, and critically, improvement in data-driven engagements for patient care and delivery of health services to last-mile communities.

## Context

Current protocols described by the National Health Mission and the Rajasthan Ministry of Health and Family Welfare outline health tracking across the maternal child health spectrum. First newlywed couples (i.e. women who may soon enter pregnancy) should be identified by the Accredited Social Health Activist (ASHA) and given a serial number known as the Eligible Couple number. During pregnancy, each mother attends a Maternal and Child Health Nutrition camp in her village (also known as the VHND camp) and receives a pictorial, paper card known as the MAMTA card as a personal record for her pregnancy and her child's upcoming infancy. The Auxiliary Nurse Midwife (ANM), who services a Subcenter Catchment area of an average 5000 individuals in plain terrain and 3000 in tough terrain, is expected to see each mother four times during pregnancy and provide antenatal care checkups (recording any signs of high risk), maternal vaccines, iron folic acid tablets, and deworming medications. During these visits, the ANM is expected to fill the mother's MAMTA card and her own Reproductive Child Health (RCH) register (a log of all patient data from the camp) with the same data. By month's end, the ANM is expected to calculate totals for key health indicators that took place at her camps in central government-mandated Forms 6-8. The ANM also turns in her linelist report from the RCH register showing the individual details of each beneficiary who attends the camps to the sector level Data Entry Operator (DEO).

The Data Entry Operator enters this linelist report into the Pregnant Woman and Child Tracking System (PCTS), a platform developed by the Government of Rajasthan. After copying values into the web portal for a given patient, the PCTS portal returns an ID for the mother or child known as the PCTS ID. These IDs, along with a due list of the next month's expected patients, are presented to the ANM by the DEO or Lady Health Visitor (LHV). The ANM is expected to write the PCTS ID on the Mother or Child's MAMTA card at the next camp checkup. ANMs are salaried government staff, but still are evaluated on their ability to reach performance targets for various health indicators such as: registrations, antenatal care checkups, and immunizations given. Primary Health Centers are mandated to have supervisory staff to conduct household "spot checks" for 10% of the beneficiary population to confirm whether mothers and children have indeed received services. Our observations from three years of field experience in rural Udaipur have shown us gaps in not only in the process of delivery of services but also in the process of data collection [2]. At the health camp level we have observed: ANMs not filling their RCH Registers with all the required 130 columns of the RCH Register, ANMs skipping sections for past obstetric history, ANMs deliberately falsifying data for blood pressure values, blood sugar vaues and urine analysis results without performing tests, ANMs filling data in personal diaries in leiu of the official register, ANMs calling migrated patients and recording follow-up status over the phone for details that require an in-person checkup, ANMs incompletely filling data on the patient's MAMTA card, particularly for immunization dates, and ANMs holding the MAMTA card until completion of the last vaccine instead of appropriately providing the card to the patient. At the primary health center level we have observed: DEOs receiving data to upload up to 30 days after the health camp, DEOs changing the dates of antenatal care checkups and dates of the patient’s last menstrual period to bypass validations, ANMs summarizing monthly indicators by hand, and ANMs manipulating data for month-end reporting for indicators uploaded into the PCTS portal. Similar findings have also been documented in a gap analysis of the PCTS system by the Columbia Earth Institute for the district of Dausa, Rajasthan [1].

## Innovation

In response to the challenges of paper-based tracking at the field level, several solutions have been developed to digitally collect MCH data in rural healthcare settings. Solutions such as eJanSwasthya, Rajsangam, Medic Mobile, Commcare, mSehat, and ANMOL share several features: a mobile application for health workers to use to digitally collect structured health data in offline settings, validations to assist health workers to take actions,  a mechanism to sync data to a cloud-based database, and a reporting mechanism such as a dashboard for district health officials to see progress against key indicators, and in some cases, use of SMS reminders for patients tracked with the system.

While these solutions have been used successfully to automate data collection, they remain vulnerable to several key limitations with respect to data accountability. Health workers may still create records offline without physically seeing beneficiaries. Furthermore, if beneficiaries move villages, which is routine in Rajasthan in the last month of pregnancy, then the new health worker often must create a duplicate record in the existing mHealth solutions and has no insight into the beneficiary’s past medical history on the digital platform. When the health worker enters the unique ID or name of the beneficiary to search the record, the beneficiary record may not be found because the data from where the beneficiary was initially registered was not synced to the current health worker’s device.

The Khushi Baby platform was designed advance the paradigm for mHealth based tracking by addressing the key issues of accountability and dependence on synchronized devices. The Khushi Baby Pendant stores the medical history of the beneficiary, allowing them to carry their updated health record to any health worker in a digital format. The health worker cannot update or create a new health record without the beneficiary’s biometric and Khushi Baby Pendant being scanned to match. The Khushi Baby platform also was designed to make culturally-informed improvements to technologigical approach, through community-inspired design of the wearable digital health record and with patient-specific, dialect-specific voice reminders for populations that are largely illiterate.

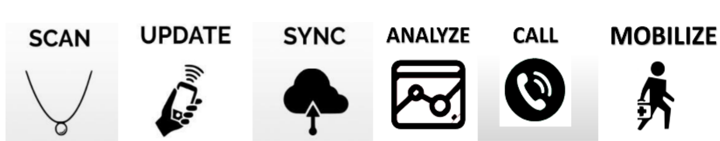
Ultimately, “breaking through stagnation for infant immunizations” will require addressing existing gaps in health system accountability, throughout the continuum of pregnancy through infant care, to drive behavior change for both patients AND health workers who make up the health system in rural India. Our undergirding hypothesis is that better data - actionable and accountable - can bring about better prevention and better care for mothers and children at the last mile.

## Intervention

Khushi Baby is a System-as-an-Intervention designed to transform MCH at the "last-mile". The system's overarching objective is to better connect health workers with beneficiaries and health officials with communities through data-driven engagements. To achieve these objectives, the KB system has multiple components, targeted for multiple stakeholders (see Appendix to review the project and intervention evolution).

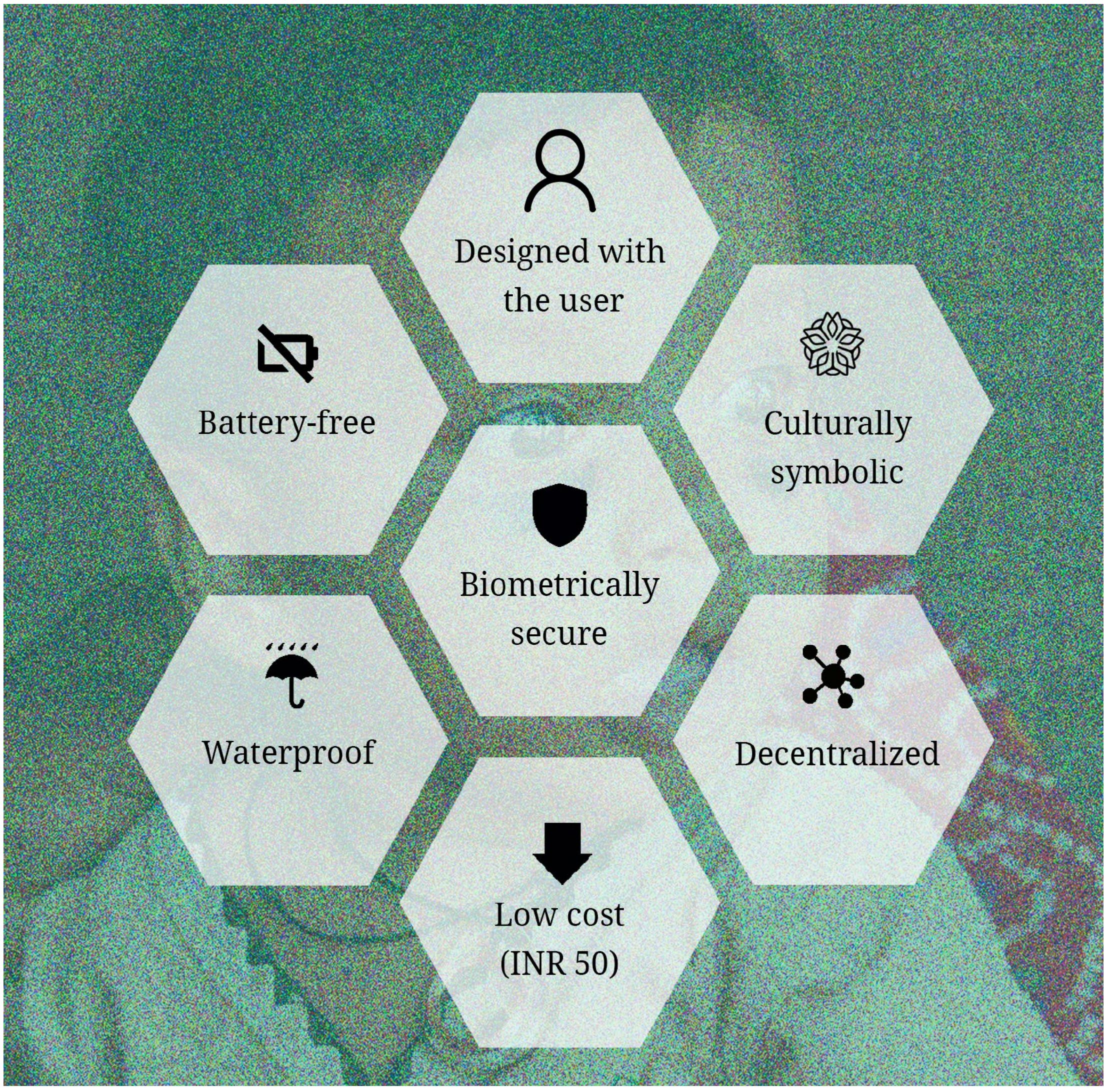
1. The Khushi Baby Pendant *for beneficiaries*
2. The Khushi Baby App *for health workers*
3. The Khushi Baby Analytics Dashboard with Automated Voice Calls *for health officials and families*
4. The Khushi Baby Personal Voice Calls *for beneficiaries*
5. The Khushi Baby WhatsApp Groups *for health worker teams*
6. Khushi Baby Field Monitors *for health workers and beneficiaries*

*Figure 1. Khushi Baby Intervention Schematic*

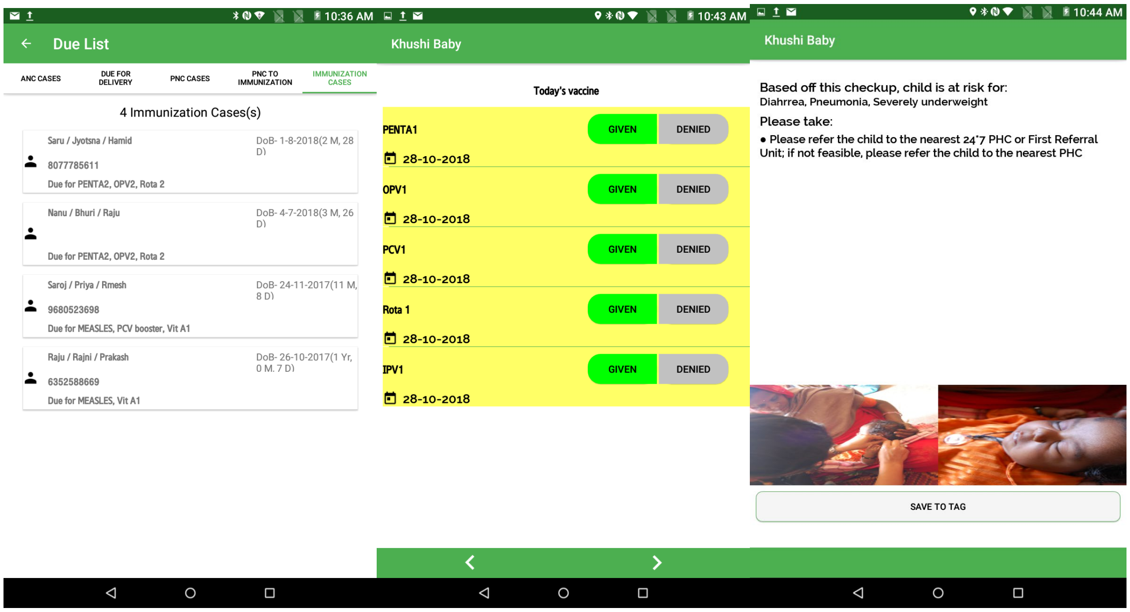




***How it works:** Mothers and children wear the Khushi Baby Pendant. Health workers (ANMs) scan the pendant with the Khushi Baby App to retrieve and systematically update health information and plan for health camps. ANMs sync data when they return to areas with connectivity for health officials to analyze reports. Automated voice calls are sent to families to remind them of upcoming camps. High risk patient reports are generated and sent via WhatsApp groups to health worker teams to mobilize on-the-ground action. Monitors support ANMs to facilitate high risk follow-up.*

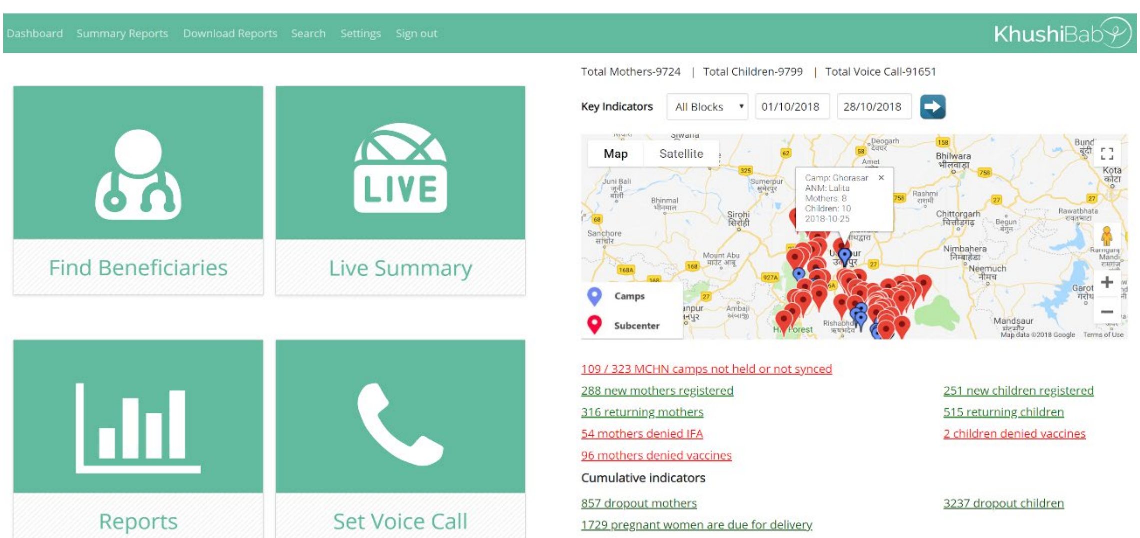
The Khushi Baby Pendant: a novel health record which uses Near Field Communication technology to store unique IDs and the full health history of the mother and child. Importantly, this pendant has been designed with a culturally-symbolic black thread, traditionally worn in these regions of India, to protect the child from evil eye or buri nazar. By tapping the pendant with the health worker's tablet, and then scanning the beneficiary's fingerprint, the health worker can access and update the patient's most recent data without the need of connectivity.




The Khushi Baby App: an Android Application (OS 4.4.3+), made for tablets, to be used by the frontline health workers (ANMs). Features include: authenticated login with biometric, health camp selection, GPS and time tracking to confirm attendance, pre-camp supply checklists, color-coded patient due lists from pregnant women to registered infants, data field validations and alerts, time taken per beneficiary tracking, check-up summary page with action items for the beneficiary, immunizations due for today logic, ability to report why vaccines, medications, and tests were not administered, camp-check out summary with updated due list, ANM summary statistics. Using the KB App, the ANM scans the beneficiary's Pendant and then their fingerprint to read and update history. If the pendant is lost, there is a mechanism to search by name for the beneficiary in the tablet to issue a fresh pendant. All data collection fields are made to be compatible with Indian National Health Mission guidelines for RCH Register Sections II and III.




The Khushi Baby Dashboard: the Khushi Baby Dashboard presents real-time updated summary statistics, targeted for health officials. Reports are available in government formats for major indicators. Specialized reports are also available for health worker attendance, supply side shortages, and high risk patients. Officials using the Dashboard have the ability to drill down by geography to the individual level (mother-child dyad) to track progress. The KB mDash is an Android Application under development designed for officials to see actionable data distilled from the Dashboard, while on-the-go from their mobile phones.



Khushi Baby Automated Voice Calls: through the KB Dashboard, officials can schedule calls for camp reminders and for maternal and child health education. These calls can be set for a specific time, geography, and for a specific beneficiary group, such as those mothers who missed their child's last immunization. Over 35, thirty-second audio clips have been recorded in the regional dialect of Mewari.

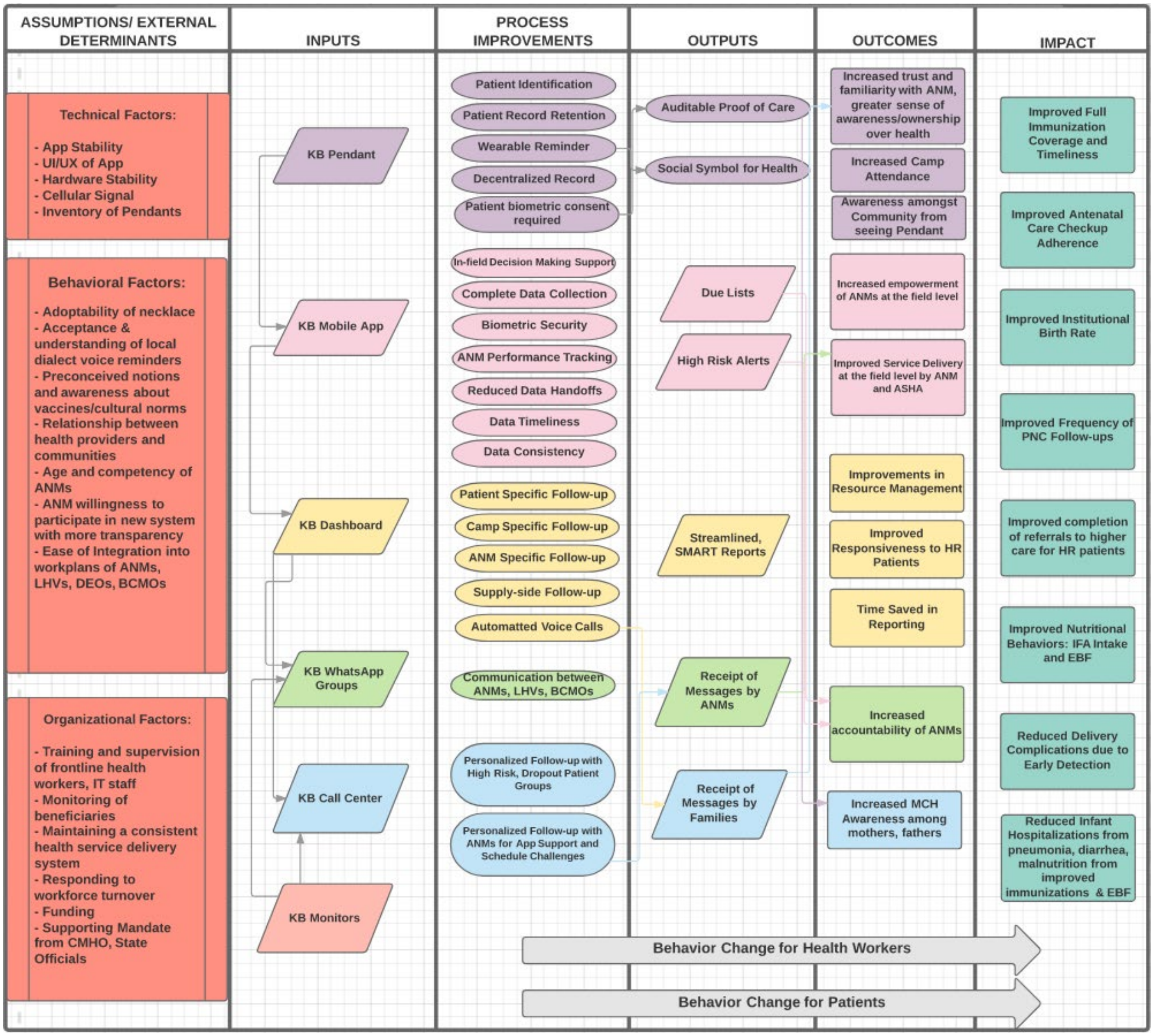
Khushi Baby Personal Voice Calls: one member from the KB Team is currently designated as the Community Engagement Expert. She calls selected beneficiary groups, such as drop out mothers, high risk mothers, and mothers in their last month of pregnancy to provide timely advice and listen to concerns. Note that this component of the intervention was not rolled out to study participants.

Khushi Baby WhatsApp Groups: ANMs and their Block Medical Officers are added to WhatsApp groups. On a weekly basis, the KB team shares high risk patient reports with the ANMs. The groups are also used to share educational content related to specific MCH themes. ANMs use the groups to report back on high risk patients from their catchment areas.

Khushi Baby Monitors: Khushi Baby field monitors are local staff who provide ANMs with training, in-field support, and assistance in following-up with high risk and drop-out beneficiaries.

## Theory of Change

*Figure 2. Theory of Change Diagram*




Khushi Baby's theory of change[9] is rooted on achieving the following specific aims:

1. To ensure an interface between frontline health worker, caregiver, and child
2. To streamline data collection on maternal child health indicators
3. To enable better planning and clinical decision making on the part of the frontline health worker
4. To improve communication between the ANM and ASHA for care coordination
5. To optimize management of limited health worker resources
6. To identify highest risk and drop out beneficiaries
7. To better educate and remind beneficiaries about importance of antenatal care and immunization
8. To change the culture of action and accountability among health workers and health officials.

[9] For full description of our Theory of Change, see our Baseline Report (p. 36, 37).

## Evaluation Framework

### Key Evaluation Questions*

1. Does the Khushi Baby System function robustly?
    a. What is the frequency of hardware and software related issues in the field?
    b. Does the KB System deliver data on time?
    c. Do beneficiaries retain pendants more than MAMTA cards?

2. Can the KB Platform improve *the quality for MCH data* in rural Udaipur when compared to status quo processes?
    a. Is the data complete for minimum, mandatory fields for infant health tracking?
    b. Is the data consistent between the beneficiary and the backend-source?
    c. Is there an impact on the checkup-related processes followed and/or the proportion of false fields entered?

3. If that the KB System works and generates higher quality data, does the KB system generate value or remain unused? Specifically how does the KB System generate differentiating data-driven engagement for better health care delivery and community engagement?

4. Does the collective KB System lead to high rates of full and timely infant immunization, and if so which factors were significant in predicting immunization outcomes? Specifically:
    a. What factors contribute to full immunization as defined by oral recall?
    b. What factors contribute to full immunization as recorded on the official health card (MAMTA)?
    c. What factors contribute to full immunization as recorded on the Khushi Baby pendant?

5. Are there any spillover effects of the Khushi Baby system on:
    a. Rate of severe acute malnutrition and moderate acute malnutrition?
    b. Infant hospitalization rate?
    c. Infant mortality rate when comparing Treatment vs. Control?

6. Comparing Treatment to Control, how have the attitudes, behaviors, and awareness for mothers and fathers with respect to the health care of their child changed?

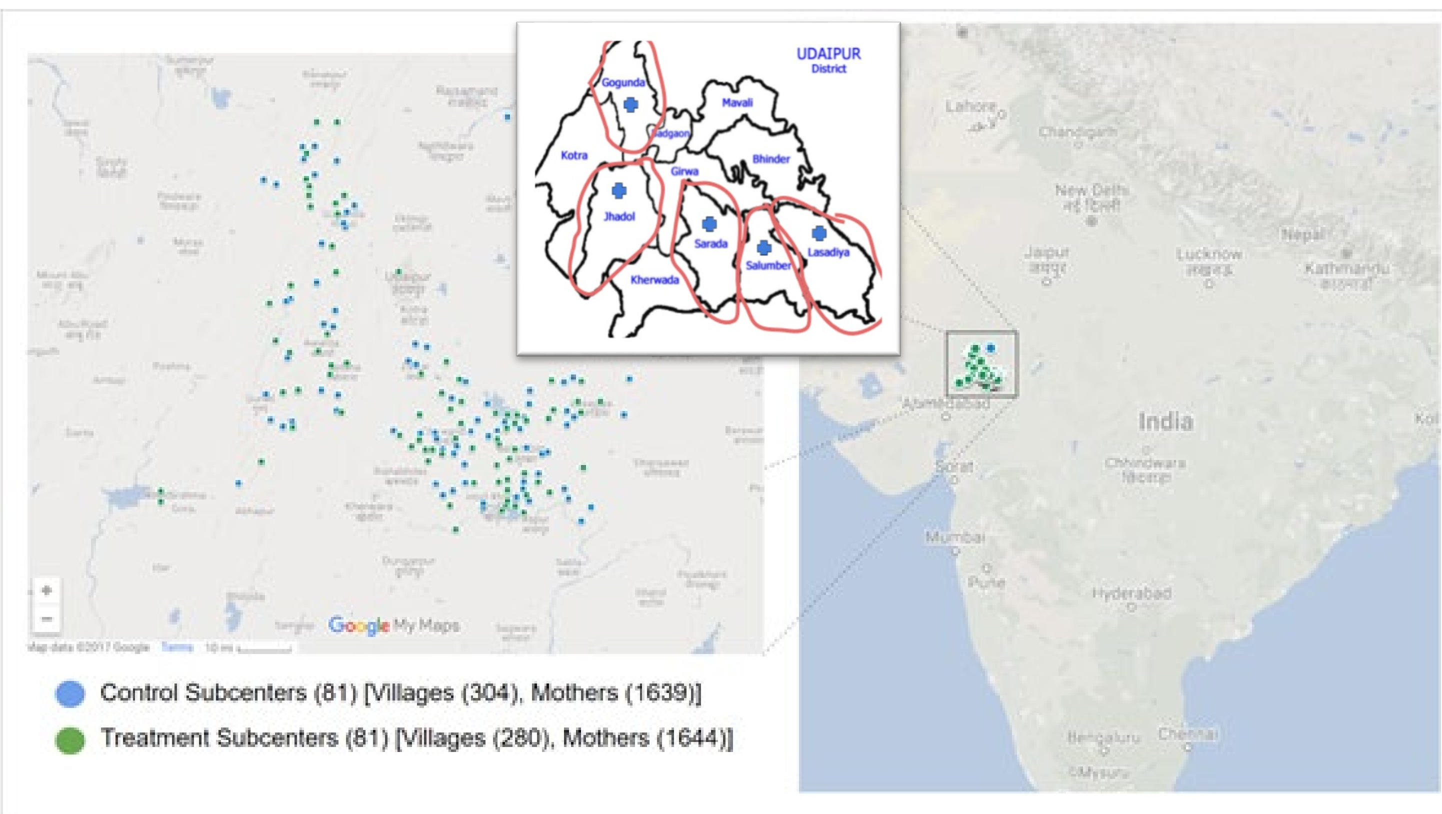
7. Comparing Treatment to Control, how have the attitudes, behaviors and practices changed for ANMs and ASHAs from Baseline to Endline?

8. What is the end user feedback on components of the system from mothers, ANMs, ASHAs, and block managers who have experienced the pendant, KB App, KB Automated Voice Calls, and KB WhatsApp groups respectively?

## Study Design*

A 162-subcenter (SCs), unblinded, cluster Randomized Controlled trial with two arms, Khushi Baby System as an Intervention and Control, took place between September 2016 and August 2018. In Udaipur there are 627 SCs, each covering a 5km radius and population of 5000 people on average, across 12 administrative blocks. SCs where earlier pilots had been rolled out in 2015-16, and SCs with other interventions such as E-Jan Swasthya, another mHealth pilot, were excluded. The SCs were further narrowed by the CMHO who allotted 5 administrative blocks from which we could select our Treatment group. Of the remaining 252 SCs, 162 were selected using a random number table. The list of 162 was approved by the CMHO in August 2016. The 162 SCs were randomized to the KB Intervention and Control groups using a stratified randomization with blocking using STATA software. Randomization was stratified on baseline full immunization coverage (FIC), as determined by PCTS for the 2015-16 Financial Year across three levels: "high performing SCs" had a reported FIC of 90%+; "mid performing SCs":60-90%; and "low performing SCs": below 60%. Randomization checks were performed to ensure balance on the stratification factor. Randomizations were run until the difference in the stratification factor was no longer statistically significant between arms. The 81 SCs randomized to Treatment were then visited by the KB monitoring team to gather ANM contact details. ANMs who were subject to retire or transition were excluded.

*Figure 3. Map of Intervention and Control Areas*




*See Appendix for Evaluation Design Framework, Changes from Pre-Analysis Plan and Monitoring Plan

Sample Requirements were calculated based on assumptions for Outcome 4A, Full and Timely Immunization as verified by oral recall, due to the relevance of the outcome towards the grant thematic window and due to the outcome having the largest projected sample size at the time of baseline for nested evaluation questions. For two study arms, an alpha of 0.05, power of 0.8, minimal detectable difference of 10% points from baseline rate of 0.5 to endline rate of 0.6 (which would be of programmatic significance), assumed loss to follow-up (LTF) of 10%, and assumed Inter-cluster Coefficient of 0.15 (high inter-cluster homogeneity), the required sample size was 165 subcenters, with 20 eligible participants from each subcenter for a total target sample of 3300. The baseline infant immunization rate was 0.241, not 0.5. Therefore, the revised required sample for follow-up through endline was 2480. The subgroup sample required for health record retention for those mothers who received a health record for their infant was 3300, using the same assumptions above and a measured baseline retention rate of 0.509. Other subgroup samples requirements were not calculated at baseline, due to limitations in the baseline coverage evaluation survey.

### *Table 1. Sample Required and Sample Achieved*

| | Sample Required for 10% Minimum Detectable Difference | Sample Enrolled | Eligible Sample followed for Outcome at Endline |
|---|---|---|---|
| Full and Timely Immunization at 12 months | 2480 (2254 without 10% LTF) | 3283 | 2243 (2254 required) |
| Health Record Retention | 3300 (3000 without 10% LTF) | 3283 | 2145 (3000 required) |

Loss-to-follow-up (LTF) reasons varied by outcome. Most households were revisited during the Endline exercise in both Treatment (81/1644 missed) and Control (69/1639 missed). LTF was due to mothers not available at the Endline visit, mothers having moved from the household, mothers who had died, mothers who had refused consent, and children who died after roll-out of the intervention. Health record analysis required receipt of a health record as a prerequisite. Mothers who did not receive the pendant or MAMTA card were excluded. Based on the endline sample achieved, health record retention as an outcome was underpowered to detect a 10% difference. Data completeness outcomes required a searchable patient ID. Mothers with duplicate Khushi Baby IDs or null PCTS IDs were excluded. Full Immunization by recall excluded children under the age of 12 or those who died before the intervention start, whereas full immunization by MAMTA card further excluded children with no corresponding MAMTA card. Children who did not meet the 6-month age criteria were excluded from malnutrition outcomes along with any who had died. Children who died before intervention start were excluded from the hospital admissions outcome and from the infant mortality outcome. For each outcome, an endline balance table was constructed to account for possible differential attrition between study arms. Differential predictors were adjusted for in the multivariate analysis where applicable. See appendix for Full Consort Diagram.

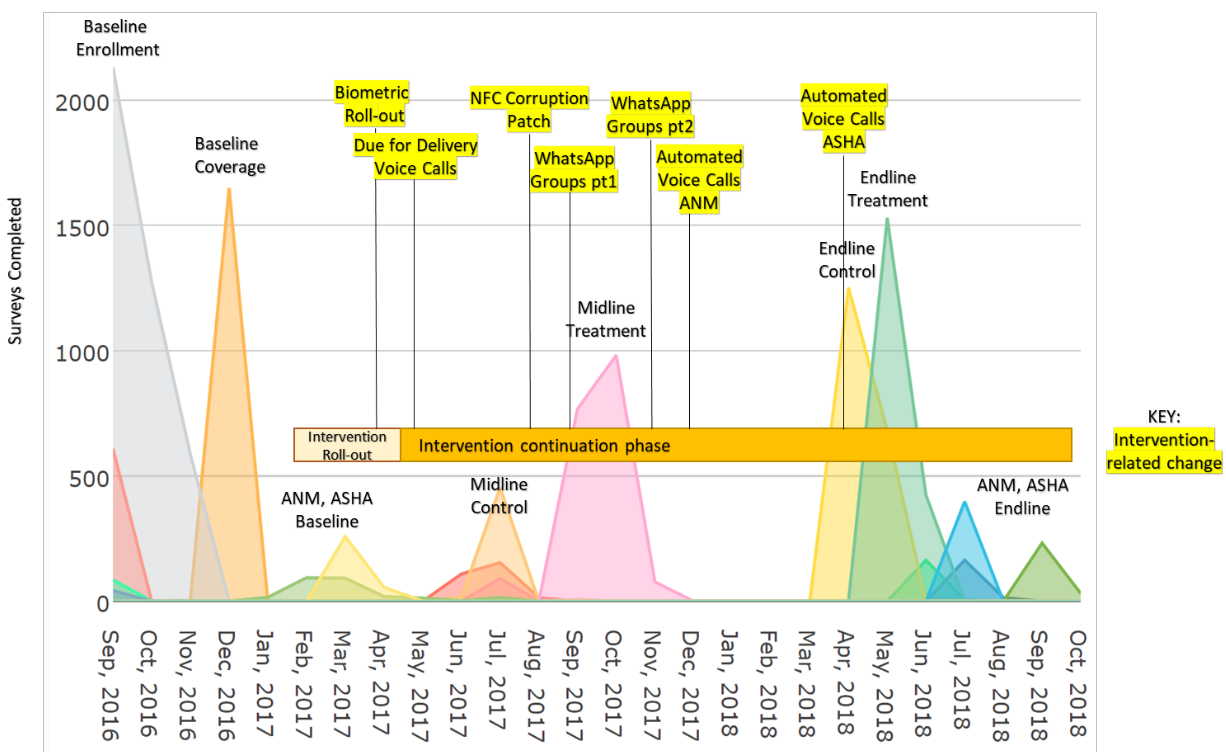
*Sampling Strategy*

*Baseline Enrollment*

The purpose of the Enrollment Survey was to select the cohort of mothers to be followed longitudinally for our impact evaluation for up to 18 months. The Enrollment Sample, derived from the Enrollment Survey was used to assess outcomes for: Data Retention, Data Completeness, Data Consistency, Full Immunization Coverage and Timeliness, Pentavalent Coverage, Severe and Moderate Acute Malnutrition, Infant Hospitalization and Mortality, and changes in maternal attitudes towards health care. The Enrollment Survey collected important baseline indicators regarding mother SES, mother MNCH awareness, and past medical history which could serve as relevant covariates in models for the stated outcomes.

For each of the 162 Control and Treatment SCs, lists of pregnant women ("linelists") were obtained from ANMs. Each mother in the linelist was visited at the household and individually invited to participate. This approach was used in order to avoid selection bias from sampling within SCs, and to ensure enrollment of pregnant women from all villages serviced by the SC. The target for enrollment per SC was 20 pregnant women. If the target could not be reached by recruitment of mothers on the linelist, mothers in the same village but not on the original linelist were invited to participate ("non-linelisters"). Mother who had already delivered were excluded. Mothers not available at the household at the time of visit had their households revisited once.

*Figure 4. Study Timeline*



*Baseline Coverage*

The purpose of the Baseline Coverage Evaluation Survey was to understand the immunization rates at baseline of the average 12-23 month old child in the control group and the treatment group subcenters for randomization stratification.

Of the 162 SCs chosen for the study, 50 Treatment and 50 Control SCs were selected using a random number generator procedure on STATA. Linelistings were obtained for all mothers with children age 12-23 months from each of the 100 selected SCs. As per 30 Cluster Sampling Guidelines from the WHO[10], a required sample size of 13 respondents were targeted per SC (level of clustering), in order to have 95% CI, 5% margin of error, with an assumed true population baseline of Full Immunization Coverage of 70%. Systematic random sampling was conducted within the SC linelistings to determine which households to visit, and the process was continued until the target number of 13 was reached in each SC. Posthoc analysis showed 12.43 respondents were surveyed per SC on average.

*Baseline ANM*

The purpose of the ANM Baseline Survey was to assess behaviors and attitudes of ANMs to the data collection process for MCH tracking, and for the routine work performed in general. Structured questionnaires were deployed to investigate: challenges, motivations, workflows, work conditions, as well as individual factors of the ANM such as familiarity with mobile phone and years spent in training. All ANMs serving the 81 Treatment and 81 Control Subcenters were selected for the ANM Stakeholder Sample (at Baseline). 166 ANMs in total were surveyed: 88 Treatment and 78 Control.

*Baseline ASHA*

The purpose of the ASHA Baseline Survey was to assess behaviors and attitudes towards MCH tracking from the perspective of the ASHA who plays a key role in recruiting mothers to the camps and in coordinating care with the ANMs. Structured questionnaires were deployed to investigate: challenges, motivations, workflows, work conditions, as well as individual factors of the ASHA. A convenience sample of 2 ASHAs was targeted for each of the 81 Treatment and 81 Control SCs. Each SC has approximately 4-5 Villages serviced, each of which has an ASHA. A total of 310 ASHAs were surveyed.

*Midline Follow-up*

The purpose of the midline survey was to follow-up with the Enrollment Sample for health card retention, data completeness, data consistency, and immunization coverage for OPV1-3 and PENTA 1-3, and to gather an intermediate time point for these indicators. All women in the Treatment group were eligible. A systematic random sample of every fourth respondent was approached in the Control group.

---

[10] World Health Organization. Vaccination Coverage Cluster Surveys: Reference Manual. https://www.who.int/immunization/monitoring_surveillance/Vaccination_coverage_cluster_survey_with_annexes.pdf

*Endline Follow-up*
The purpose of the endline survey was to gather data on all evaluation outcomes on the Enrollment Sample. Additionally, any variables missed during the baseilne collection were also collected during the endline survey, such as distance to the health camp. All woman originally enrolled were eligible for this survey.

*Endline ANM Survey*
The purpose of this survey was to gather feedback from the ANMs on Outcome 7. Topics addressed included comfort with and performance of the Khushi Baby system, confidence with conducting essential duties in the setting of the KB system, and impact of the KB system on camp turn up. All ANMs enrolled at baseline were eligible for this survey. 147 ANMs were surveyed in total.

*Endline Camp Observation*
The purpose of this survey was to observe the essential checkup processes and data entry processes performed by the ANM at the camp for Outcome 2C - Data Validity. Checkup processes for new and returning mother and child were observed along with subsequent data entry steps. All 147 ANMs at endline were eligible for this camp observation, but observation was completed before the full sample could be observed. However, the order of ANM observation was randomized. During each camp observation session, KB monitors were given a target of observing 1 new mother, 1 returning mother, 1 new child and 1 returning child. In Control SCs, 31 ANMs who performed 32 mother checkups and 35 child checkups were observed. In Treatment SCs, 59 ANMs who performed 68 mother checkups and 75 child checkups were observed.

*Endline Focus Groups - Mothers*
The purpose of this survey method was to gain a qualitative understanding of beneficiary feedback towards various components of the KB System such as the pendant and voice calls and their suggestions for how to improve these components moving forward. One facilitator and two observers were present to probe discussion topics and to capture verbal and non-verbal responses of the group respectively. For each of the five geographical blocks, a target was set for two to three focus groups at the village level. These villages were selected using a convenience sample accounting for geographical access, availability of the ASHA, and availability of the beneficiaries who had received the KB Pendant. The survey was notably conducted during the maize harvest season thereby reducing the availability of a certain set of beneficiaries. 92 Mothers were included in a total of 11 focus groups, with a minimum of two focus groups from each geographical block.

*Endline Key Informant Interviews*
The purpose of this set of interviews was to gain an understanding of the perspectives of LHVs, BPMs, BCMOs, MOICs, and the CMHO, who oversaw theTreatment SCs, on the impact of the Khushi Baby system, their individual interaction with the system to date, and their ideas for system improvement moving forward. A total of 36 LHVs, 36 MOICs, 5 BPMs, 5 BCMOs, and 1 CMHO were concerned with the Treatment area. A convenience sample was selected with attempts to

cover each block and reach health officers who were both geographically proximal and distal from the Udaipur HQ. The target sample was 3 LHVs and MOICs from each block, and all 5 BPMs, BCMOs, and the CMHO for a total sample of 41. In total, 27 interviews were conducted with 10 LHVs, 8 MOICs, 3 BCMOs, 5 BPMs, and the CMHO.

*Data Collection and Management*

Data for all surveys except the endline focus group discussions and key informant interviews were collected on the SurveyCTO mobile application by a team of Khushi Baby Field Surveyors, hired as temporary contractors, and given an initial 3-day workshop with regularly weekly meetings at the Khushi Baby HQ. The Khushi Baby Core Team would evaluate the data collected daily and examine data for duplicates and missing mandatory fields. During the endline follow-up survey, it was noted that MUAC values were being reported with rounding bias on several hundred infants. Surveyors were informed to round to the nearest tenth as per the tape and distribution of values improved over the next month and a half. Backchecks and spot checks were performed by Khushi Baby Core Team members after reviewing the weekly data. Data were merged using unique identifiers for respondents and subcenters.

## Methods for Analysis

### *Approach to Quantitative Analysis:*

Descriptive analyses were performed on all Outcomes 1-8 for frequencies and distributions. Differences between Treatment and Control groups were calculated on Outcomes 1C, 2A, 2B, 2C, 4A, 4B, 4D, 5A, 5B, 5C, and 7 using parametric two-tailed, two-sample z-tests for proportions on an *Intent-to-Treat* basis. For Outcome 6 we compared distributions using a non-parametric, Chi-square goodness-of-fit test between Treatment and Control groups and made individual comparisons using appropriate parametric methods mentioned above.

For outcomes 1C, 4A, 4B, 4D, 5Ai, 5Aii, 5B, 5C, those outcomes which had a significant difference between study groups were considered for regression analysis to adjust for potential confounders. Outcomes 4A, 4B, 4D, and 5Aii were considered for regression. Additionally, the treatment-specific Outcome 4C was regressed to corroborate the results from the analyses of 4A and 4B. Logistic regression was used on these binary outcomes, using their respective intent-to-treat sample, according to the following formulation:

$$Logit\left(Y_{ijkl,post}\right) = \alpha + \beta T_{kl} + \gamma X_{ijk} + \delta Z_{jk} + \rho W_k + \mu_{jk} + \varepsilon_{ijk}$$

In this equation, $Y_{ijkl,post}$ represents the binary outcome at endline for the woman $i$ in village $j$ in subcentre $k$ in block $l$. $T_{kl}$ is a binary vector for subcentre $k$ being in the treatment group. $X_{ijk}$, $Z_{jk}$, and $W_k$, respectively represent the vectors of individual-level, subcentre-level, and administrative block-level covariates (only one categorical covariate for administrative block as a fixed effect). The subcentre-level error term is represented by $\mu_{kl}$, and the individual level error term is represented by $\varepsilon_{ijkl}$. Individual-level covariates included variables such as age, marital status, education, caste, socioeconomic status index, a maternal and child awareness index score, access to a mobile phone, distance to the health camp, child sex, child illness episodes, restrictions imposed by family members, frequency of ASHA visits, time spent waiting at camps, and number of intervention-related calls received. Subcenter-level covariates were specific to the ANM, including (but not limited to) factors such as ANM age, education, distance to furthest camp, frequency of visit to the Primary Health Center, total requirement of reporting burden, and access to mobile phone. Subcenter level covariates also included number of external overlapping interventions, such as the Mission indradhanush campaigns for immunization. $\beta, \gamma, \delta, \rho$ represented coefficients for study group, individual-level, subcentre-level, and block-level covariates, with $\beta$ as the coefficient of interest, which represents the increase or decrease in the log-odds of the outcome for an average woman in a subcentre randomized to receive the intervention.

To build a parsimonious model, we had to select from over 260 measured variables related to the mother, the ANM, the ASHA, the geography, and the components of intervention received. Our target model intended to include the 15 most significant covariates associated with the outcome.

The systematic approach to variable selection for the regression model was as follows:

1. List variables to be associated with the outcome based off theory and prior literature as the “base model”
2. Impute missing data using predictive mean matching for continuous and binary variables and mode substitution for categorical non-binary variables with the MICE package in R
3. Reduce dimensionality via indices for socioeconomic status and health awareness (see Appendix)
4. Assess differential distribution of measured variables between Treatment and Control groups at Endline
5. Assess the unadjusted, univariate association between each variable and the outcome
6. Perform principal components analysis to further reduce dimensionality according to the methods described by Zhang et al.[4]
7. Assess which top three principle components explained the largest degree of variation in the sample
8. Assess which remaining principal components were most delineated by selecting those principle components which had at least one major loading with a covariate of |0.18|
9. Assess which principle components were most statistically significantly associated with the outcome using a multivariate model built empirically, using forward selection
10. Variables from steps 1, 7, 8, 4, and 5 were then sequentially added, with non-significant variables dropped (forward selection), to construct the final model, using the Likelihood Ratio Test to assess difference between sequential models, and the objective AIC criterion to determine the overall best fit model under the conditions for a parsimonious model, as described in [5]. The GLMER package in R was used for this step to specifically adjust for random effects of the cluster level and to check for highly multicollinear variables. For details on this statistical package implementation, see the following reference: [6].

Variables were included as categorical or continuous based on the original method of measurement unless there was a literature-based rationale for converting continuous variables into specific categorical bins. Continuous base model variables found to be non-significant in the final step were considered for categorical binning and retested against the model, if there was a non-linear trend seen between the variable and the outcome. Of note, treatment-specific Outcome 4C followed the same procedure above with exception of step 4, which was not applicable, and for step 10, GLM (not GLMER) was used. Both significant covariates from the final model and non-significant covariates from the base model are interpreted in *Section 5: Findings*.

### *Approach to Qualitative Analysis*

Individual quotes spoken in Hindi or Mewari by the mothers in the focus groups and Hindi or English were recorded in Hindi or English by the Khushi Baby Core team. Quotes were grouped according to each open-ended question-stem topic. Representative quotes were then selected for variety of role, level of detail, variety of opinion, and uniqueness of response. Word and phrase frequencies were not quantitatively analysed. Key thematic areas from the full set of quotes and representative set of quotes were extracted and interpreted.

*Approach to Addressing Sources of Bias:*

From a quantitative standpoint, our objective was to provide an unbiased estimate of the effect of the Khushi Baby intervention on full infant immunization, among other outcomes related to both data and health of the child. Several sources of bias were present in the study. First, the Khushi Baby team was responsible for designing the study, modifying the intervention, analysing the data, and interpreting the results which portends evaluator bias. An external organization may have been able to conduct an impact evaluation with less of a vested interest in the role of the intervention on the outcome. Logistically and financially, this option was not feasible. Moreover, the degree to which an external agency may have been able to evaluate the project would likely be limited in comparison. Key to reducing bias was making available the Pre-analysis Plan and Baseline Report and justifying any deviations in intended analysis. Also, data collection was conducted by field monitors and surveyors hired as consultants through an external agency, A to Y Manpower, and detailed code and anonymized data will be made available through the supplement for others to replicate our analysis.

Sampling bias was addressed through the study design. Women chosen for the study were selected at the household level, not at the health camp level, and from an ANM linelist of each village. There was a possibility that the ANM and ASHA-collected censuses of reproductive age women may have left out migratory populations or members from lower caste groups, but our analyses found no differences in SES score, caste, or outcomes between linelisters and non-linelisters (see Appendix for details).

With respect to the data collected, several forms of bias may have contributed including recall bias especially with respect to Outcome 4A, and that too among a population already established to have a low baseline health literacy in the Baseline Coverage Evaluation. To reduce recall bias, the survey questionnaire was designed to indicate the injection site and route of administration, using culturally relevant cues for the various vaccines modified from the National Family Health Survey.

With respect to Outcome 4B, misclassification bias was a concern, given that ANMs were found to keep MAMTA cards to themselves, not always record the given status of the vaccine, or in some cases, record the given status without the date received. The effect of such misclassification would likely result in a lower proportion of respondents completing the full immunization requirements. At the same time health incentives for ANMs to submit MAMTA cards showing full immunization may have driven misclassification bias to overestimate the true coverage outcome. We expected the direction of misclassification to be equivalent in study groups given the randomized design. Outcomes 4A and 4B were included in the study design to corroborate findings from both approaches to measure full immunization, and in doing so account for each method's own measurement bias.

Hawthorne effects from the Khushi Baby surveyors on the individual women enrolled were unlikely to have played a factor in the observed differences in outcomes. Women were approached and asked about the outcome at most three times: baseline, midline, and endline. The effect of observation almost certainly played a role on ANMs, exclusively in the Treatment subcenters, who received varying degrees of interaction and supervision from Khushi Baby field monitors. It is important note however, that outcomes were specific to the individual level and Khushi Baby monitors were part of the intended intervention being evaluated for scale-up. John Henry effects of individuals in Control groups, overcoming a known differential application of the intervention, were minimized by the clustered design, which ensured that individuals of a same subcentre (5km radius) would be relatively blinded to interventions in subcenters outside of the normal perimeter of their daily activities and less likely to cross-over to the intervention group. Contamination of subcentres due to overlapping interventions was accounted for both in the randomized design and in the analysis phase, with specific attention to the Mission Indradhanush door-to-door vaccination campaigns that took place in select subcenters.

From a qualitative standpoint, the objective was to present a representative sample of impressions (positive, negative, and group-supported) of the intervention from various stakeholders and beneficiaries. In this case, Khushi Baby Core Team members were involved in these focus group discussions and in depth stakeholder interviews. The room for evaluator bias was notable here given that the Khushi Baby team was involved with the data collection process. To reduce bias in focus group discussions, a female member (who regularly calls high risk mothers for personalized reminders) was selected as the facilitator, with two other members as silent observers. The facilitator would engage the women in ice-breaking activities to gain comfort and ask open-ended questions. When certain voices began to dominate the conversation, the facilitator would intervene to ask quieter members to also share their thoughts. The facilitator would also probe for women to share thoughts in more detail. In both focus group discussions and in stakeholder interviews, respondents were explicitly asked to suggest areas of concern and improvement to balance against any tendency to appease the evaluator.

### Ethics and Transparency

The study was approved by a local ethics board at the Center for Operation Research and Training in Vadodara, Gujarat in July 2016 prior to the onset of the trial. Key ethical issues addressed included data privacy, extra burden placed on health workers, administrative burdens placed by randomization, time required from survey participants who came from vulnerable populations, and concerns regarding safety of the necklace. Oral consent was taken from participants at the beginning of each survey process. Written consent was also taken for participants in the baseline enrollment and coverage surveys. The study was registered online at clinicaltrials.gov with the following protocol ID: TW10.1078. Labeled anonymyzed datasets (without audio files) and statistical analysis code in R and Python with documentation can be found in the Appendix.

## Findings

### Implementation Fidelity

The initial target for implementation uptake was for 81 ANMs to use the KB App, 5 Block Medical Officers and 5 Block Program Managers to use the KB Dashboard, 38 LHVs and DEOs to use the KB Dashboard, and for the CMHO to use the KB Dashboard. We projected 10,000 beneficiaries to receive Khushi Baby pendants in total over the course of the evaluation time frame. The actual implementation uptake from February 2017 to June 2018 (the evaluation period) showed that the KB App was used by 87 ANMs, although significant challenges were faced. Although trainings were provided to LHVs (90), Data Entry Operators (28), and Medical Officers (42), these health workers and officials did not consistently use the KB Dashboard over the evaluation period. Deviations in uptake from health officials may be principally attributed to the fact that our system did not collect data from all beneficiaries in the respective catchment areas (the block level) of these health officials due to the nature of our randomization (at the subcenter level). Access to computers was not a barrier for many of these health workers, but in many cases, computers of these health workers and officials were found to be largely unused (except for the case of data entry operators). By the end of the evaluation in 2018, 20,000 beneficiaries had received KB pendants.

We faced challenges increasing uptake of the mobile app among ANMs especially between February and November 2017 for several reasons. First, many of the tablets used in the field had hardware failures. Second, some proportion of the replaced tablets, due to an error in the software library, caused several hundred records to be corrupted on the KB Pendant and KB Backend. These technical issues and other weekly crashes within the app directly caused many ANMs to express frustration and desire to abandon the KB App. In particular, we observed that ANMs were less likely to use the KB App without direct supervision from the KB Monitor. This was particularly notable by examining the synced data when field monitors were cross-covering for field surveyors during months with survey exercises and during summer months when heat would limit monitor travel. Of note, ANMs in the Lasadiya Block were particularly resistant to using the system despite over 10 visits to meet with them and their supervising medical officers. In Lasadiya, the Chief Medical Health Officer did not have political power to enforce certain mandates due to the Medical Officer also holding the political post as a state representative.

Nine hundred ninety-four mothers received the Khushi Baby Pendant. Two hundred seventy-one mothers did not receive the Pendant, of whom 23 never attended a health camp. This may be due to ANMs not using the app for each patient who visited the camp. Four hundred eighty-two mothers claimed to have received the KB Voice call with 328 of the mothers with confirmed receipt of the Khushi Baby voice call via the Khushi Baby backend out of a denominator of 1189 mothers who had children alive during the intervention. During baseline 83.9% of mothers reported having access to a phone and 29.9% reported owning their own phone. The proportion of mothers who received calls related to their infant's immunization was closer to the proportion of mothers who owned their own phone at 27.6%. The difference between the backend verified confirmed calls and recall-reported received call counts could be attributed to recall bias on the part of the mothers. Field reports suggested that initially mothers were giving their husband's phone number (or the father's phone number) before more recently starting to give their own phone numbers.

Overall major barriers to implementation fidelity included technical and organizational factors, as anticipated by the theory of change. Financial strain played a significant role. For a period of 3 months, field staff salaries had to be delayed due to a delay in two tranche disbursements of two separate grants. Limited staffing at the head quarter permitted technical hardware bugs to go unnoticed for two months before detection. One could conclude that significant barriers were overcome through this process, although the challenges experienced in this first-time implementation likely limited the full potential for impact.

## Khushi Baby System Functionality

*Table 2. Khushi Baby Platform Performance Indicators*

| Key Performance Indicator | Value | Data Source |
| --- | --- | --- |
| ANMs who needed a tablet replacement | 87 | Field Monitors |
| Pendants replaced due to corruption from software | 200 (Approximate) out of 20,000 | Khushi Baby Backend |
| Average weekly KB App crashes | 35 (Approximate) | Fabric.io Backend |
| Median time to sync data (hours) | Midline: 4.0<br>Endline: 3.75 | Khushi Baby Backend |

Performance Indicators were poor with respect to hardware stability and software stability especially in early stages when all tablets had to effectively be replaced. Nearly 200 app related issues were identified and closed, yet remaining weekly crashes are exclusively related to ongoing issues with custom software for handling the Near Field Communication module for a customized Data Mini tablet which was sent as a replacement. Other common issues reported by ANMs included crashes due to negative use cases in navigating between certain pages with complex page logic. Performance indicators were positive for the time to receive data synced from field after completion of camps with a median time of less than four hours per record from the time of creation or update at the health camp. With 93-98% crash-free sessions during Endline months, app stability still has some room for improvement.

*Figure 5. Khushi Baby App Crash Frequency*


Midline retention of the health record in the Treatment group was significantly higher at the $p < 0.01$ level. Proportion of health cards (pendants) retained in Treatment was 82.4% (713/865) and 74.8% in the Control (240/321). At Endline, retention of the health record in the Treatment group dropped to 78.3% (778/994) whereas for Control, retention increased to 77.9% (896/1150), which may be attributed to covering a larger sample than midline. The difference at Endline was not significant (p=0.840).

*Table 3. Patient Health Record Retention at Midline and Endline*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| | **Records Retained** | **Total Received** | **proportion** | **Records Retained** | **Total Received** | **proportion** | **p-value** |
| Midline | 713 | 865 | 0.824 | 240 | 321 | 0.748 | 0.0031** |
| Endline | 778 | 994 | 0.783 | 896 | 1150 | 0.779 | 0.840 |

## Khushi Baby System Data Quality

*Table 4. Data Quality in Treatment vs. Control Subcenters*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| | **Records Correct** | **Total** | **proportion** | **Records Correct** | **Total** | **proportion** | **p-value** |
| Data Completeness | 5375 | 5446 | 0.9870 | 4583 | 5936 | 0.7721 | <0.00001 |
| Data Consistency | 761 | 776 | 0.981 | 425 | 545 | 0.78 | <0.00001 |

Out of a total of 5446 mandatory fields for infant registration[11] for patients who presented pendants, 5375 were non-null. The average data completeness for a mother in treatment was 21.6% higher than that for a mother randomized to control, without adjustment for confounders

[11] Mandatory fields include: mother name, child date of birth, child weight at birth, child sex, date of registration, BCG status, Penta

and this difference was statistically significant at the p < 0.001 level. Out of a total 776 infants with a Khushi Baby pendant, 761 pendant IDs (98.1%) were successfully matched on the Khushi Baby backend compared to 81.0% matched between MAMTA card PCTS ID and PCTS backend. On average, data consistency was 17.1% higher in Treatment without adjustment for confounders, and this difference was statistically significant at the p < 0.001 level.

*Table 5. Data Entry steps on MAMTA cards in Treatment vs. Control subcenters*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| Data Validity | **Steps** | **Total Due** | **proportion** | **Steps** | **Total Due** | **proportion** | **p-value** |
| Infants | 132 | 504 | 0.262 | 66 | 245 | 0.27 | 0.830 |
| Pregnant Women | 351 | 707 | 0.496 | 150 | 301 | 0.498 | 0.960 |

*Table 6. Data Collection and Checkup Processes Followed by Study Group*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| Checkup Processes | **Processes** | **Total Due** | **Proportion** | **Processes** | **Total Due** | **Proportion** | **p-value** |
| Infants | 703 | 954 | 0.737 | 346 | 466 | 0.742 | 0.820 |
| Pregnant Women | 722 | 1059 | 0.682 | 321 | 426 | 0.754 | 0.00623*** |

Data validity was measured as the percent of correct data entry steps onto the MAMTA card out of total required data entry steps for the checkup and confirmed by direct field observation of ANMs. There was no statistically significant difference in proportion of correct MAMTA card data entry in Treatment vs. Control groups. Notably correct data entry in Pendant vs. MAMTA card was not compared here because in order to save data onto the pendant the majority of fields were required to have entry above. Data collection and standard care processes were also recorded by field monitors. There was no difference in child processes followed between ANMs in Treatment and Control subcenters, however, ANMs in Control subcenters were found to have a higher completion of required antenatal care data collection and checkup processes. This difference was statistically significant at the p < 0.001 level. There was no statistically significant distribution in new children and returning children for which checkups were observed in the Treatment and Control groups. There was a statistically significant different distribution of new mother and returning mothers between study groups with more returning mothers observed for Control ANMs, with a higher proportion of correct processes for returning mothers and common processes for new and returning mothers when in Control when compared against Treatment.

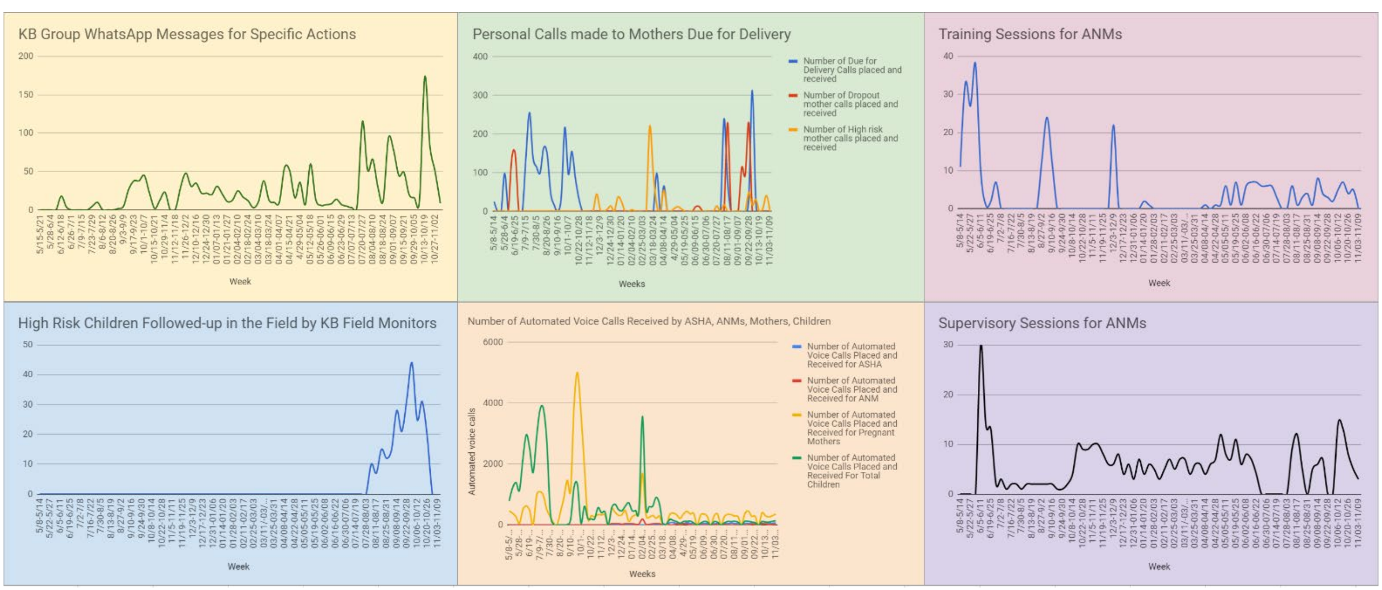
## Khushi Baby System Data for Action

*Table 7. Key Performance Indicators for Data Use for Action*

| Indicator | Average per week | Total (May 2017 – Nov 2018) |
|---|---|---|
| WhatsApp Group Messages for action | 23.7 | 1876 |
| Trainings given to ANMs, LHVs, BCMOs | 7.3 | 574 |
| Automated Voice Call Reminders Received by Pregnant Women | 539.8 | 42468 |
| Automated Voice Call Reminders Received by Mothers of Newborns | 611.3 | 48296 |
| Calls to an ANM placed by the BCMO/LHV for equipment issues | 0.7 | 54 |
| Calls to an ANM placed by the BCMO/LHV for Medication or Vaccination Stock-outs | 0.5 | 42 |
| Calls to an ANM placed by BCMO/LHV for High Risk Mothers | 0.5 | 39 |
| Calls to an ANM placed by a BCMO/LHV for High Risk Children | 1.0 | 78 |
| KB Call Center calls received by mothers due for delivery | 34.5 | 2728 |
| KB Call Center calls received by High Risk mothers or mothers with high risk children's | 9.2 | 730 |
| KB Call Center calls received by Dropout Mothers or mothers with children who are dropouts | 13.2 | 1050 |
| Supervisory Sessions held at low performing camps | 5.3 | 416 |
| Focus Groups with low performing ANMs | 0.3 | 27 |

The above key performance indicators for data for action come from a health worker team comprising of a total of 36 LHVs, 36 MOICs, 5 BPMs, 5 BCMOs, and 1 CMHO, between 81 to 87 ANMs, and approximately 350 ASHAs in the Treatment area. Most engagement came from automated voice call reminders through the KB dashboard, WhatsApp group messages exchanged by heatlh worker teams, and personalized voice calls to high risk and drop out mothers from the KB Call Center. More recently there has been an increase in weekly WhatsApp engagement and high-risk follow-up in the field by KB Monitors.

*Figure 6. Time Series Progression of Key Performance Indicators for Data Driven Action*



## Khushi Baby System Effect on Infant Immunization

### *a) Baseline Coverage Results*

The baseline coverage survey included 1243 participants with children between 12-23 months of age. A balance table was constructed to assess for differences in baseline complete immunization rates between treatment and control centers, along with other relevant predictors. The results of this balance table can be found in Tables 10a and 10b below. Predictors considered included the broad categories of demographics, antenatal care factors, delivery factors, and infancy & immunization care factors. Specific indicators were: survey block distribution, mother's age, mother's education, mother's caste category, father's age, father's education, who was the primary caregiver of the child, whether the mother ever attended an antenatal care checkup, the number of ANC visits completed, which locations did the mother receive ANCs, the distance to the camp, the time to reach the camp, the time waiting at the camp, the time spent with the ANM during examination, whether there was someone to remind the mother for ANCs, the frequency of ASHA household visits during pregnancy, which people and which formats of reminders were received by the mother during pregnancy, if the mother was ever prevented from attending health camps by family members, whether the mother received an immunization card, with what frequency checkup data was recorded on the card, whether the mother received a TT shot, whether the mother received iron folic acid tablets, whether the mother experienced hunger during pregnancy, whether the ANM did not answer questions during a checkup, the location of the child's birth, who assisted with the delivery, whether colostrum was given to the child, whether there was a post-delivery hospitalization of the mother or child, whether whether incentives were received after delivery, the duration of breastfeeding exclusively, the immunizations received by the child, and the reminders for immunization and their frequency.

*Table 8a. Baseline Coverage Balance Table*[12]

| | Control | Treatment | p | test |
|---|---|---|---|---|
| n | 612 | 631 | | |
| **Demographics** | | | | |
| Mother's age (median [IQR]) | 25.00 [23.00, 28.00] | 25.00 [23.00, 28.00] | 0.322 | nonnorm |
| Mother's education (%) | | | 0.0246 | |
| Never went to school | 350 (57.2) | 339 (53.7) | | |
| Grades 1-5 | 98 (16.0) | 127 (20.1) | | |
| Grades 6-8 | 86 (14.1) | 88 (13.9) | | |
| Grades 9-10 | 39 (6.4) | 40 (6.3) | | |
| Grades 10-12 | 29 (4.7) | 18 (2.9) | | |
| Graduate | 9 (1.5) | 7 (1.1) | | |
| Post-Graduate | 1 (0.2) | 6 (1.0) | | |

[12] We used the tableone library for R which automates comparisons between many comparisons between secondary exposure variables, depending on the variable type and distribution.[12] Categorical variables were compared with Chi-square test, continuous variables normally distributed were compared with one-way ANOVA (i.e. student's T test), and non-normal continuous variables compared with a Wilcoxon Rank Test.

## *Table 9b. Baseline Coverage Balance Table (Continued)*

| | **Control** | **Treatment** | **p** | **test** |
|---|---|---|---|---|
| Mother's caste category (%) | | | 0.0150 | |
| General | 49 (8.0) | 51 (8.1) | | |
| OBC | 46 (7.5) | 54 (8.6) | | |
| SC | 23 (3.8) | 49 (7.8) | | |
| ST | 258 (42.2) | 262 (41.5) | | |
| Distance to health camp (%) | | | 0.484 | |
| 0-0.5 km | 194 (32.4) | 167 (26.9) | | |
| 0.5-1 km | 213 (35.6) | 237 (38.2) | | |
| 1-1.5 km | 100 (16.7) | 123 (19.8) | | |
| 2-4 km | 38 (6.3) | 39 (6.3) | | |
| 4-6 km | 5 (0.8) | 5 (0.8) | | |
| **Antenatal Care History** | | | | |
| ASHA visit frequency in weeks during pregnancy (median [IQR]) | 3.00 [2.00,4.00] | 3.00 [2.00,4.00] | 0.0298 | nnorm |
| Any digital/voice-based reminder for ANC = Yes (%) | 44 (7.3) | 37 (6.0) | 0.619 | |
| Ever received MAMTA card = Yes (%) | 289 (92.9) | 270 (90.0) | 0.250 | |
| MAMTA card present at Baseline (%) | | | 0.317 | |
| No | 310 (50.7) | 300 (47.5) | | |
| Yes | 301 (49.2) | 331 (52.5) | | |
| At least 4 ANC visits completed (%) | 90 (14.7) | 106 (16.8) | 0.530 | |
| Ever received TT vaccine (%) | | | 0.994 | |
| No | 18 (3.0) | 19 (3.1) | | |
| Yes | 560 (94.6) | 575 (94.6) | | |
| Ever received iron folic acid tablets = Yes (%) | 566 (95.6) | 579 (95.2) | 0.861 | |
| Ever hungry during pregnancy = Yes (%) | 24 (4.0) | 31 (5.0) | 0.486 | |
| **Delivery History** | | | | |
| Child birth location (%) | | | 0.0373 | |
| Home | 78 (13.0) | 92 (14.8) | | |
| In transit (to hospital) | 1 (0.2) | 3 (0.5) | | |
| Sub center | 12 (2.0) | 6 (1.0) | | |
| PHC | 163 (27.2) | 128 (20.6) | | |
| CHC | 257 (42.9) | 277 (44.7) | | |
| Private Clinic | 23 (3.8) | 29 (4.7) | | |
| Other (specify) | 4 (0.7) | 1 (0.2) | | |
| Child Gender = Male (%) | 305 (49.8) | 309 (49.0) | 0.804 | |
| Pregnancy complications requiring hospital visit = Yes (%) | 97 (16.2) | 104 (16.8) | 0.845 | |
| Any post-birth hospital visits for the infant = Yes (%) | 67 (11.2) | 90 (14.5) | 0.0989 | |
| ASHA Home visits post pregnancy (median [IQR]) | 2.00 [0.00,2.00] | 1.00 [0.00,2.00] | 0.00467 | |
| Colastrum given after birth = Yes (%) | 543 (90.7) | 548 (88.4) | 0.232 | |
| **Infant Care & Immunization History** | | | | |
| Exclusive breastfeeding duration (%) | | | 0.210 | |
| Till end of first month | 14 (2.3) | 16 (2.6) | | |
| Till end of second month | 9 (1.5) | 13 (2.1) | | |
| Till end of third month | 27 (4.5) | 21 (3.4) | | |
| Till end of fourth month | 53 (8.8) | 79 (12.7) | | |
| Till end of fifth month | 143 (23.9) | 132 (21.3) | | |
| Six or more months | 310 (51.8) | 330 (53.2) | | |
| Immunization_reminder_method_Family_Member (%) | 35 (6.1) | 37 (6.2) | 1 | |
| Immunization_reminder_method_Any_other_health_worker (%) | 19 (3.3) | 17 (2.9) | 0.793 | |
| Immunization_reminder_method_ASHA_came_to_home (%) | 512 (88.7) | 526 (88.4) | 0.931 | |
| Immunization_reminder_method_ANM_or_by_ASHA (%) | 36 (6.2) | 43 (7.2) | 0.577 | |
| Immunization_reminder_method_SMS_of_phone (%) | 13 (2.3) | 24 (4.0) | 0.115 | |
| Immunization_reminder_timings_Everytime_before_ camp (%) | 547 (96.0) | 558 (95.4) | 0.734 | |
| Immunization_reminder_timings_Sometimes (%) | 24 (4.2) | 25 (4.3) | 1 | |
| Immunization_reminder_timings_After_missing_immunization (%) | 1 (0.2) | 2 (0.3) | 1 | |
| Polio_count (median [IQR]) | 3.00 [2.00, 3.00] | 3.00 [2.00, 3.00] | 0.365 | nnorm |
| BCG_Administration (%) | | | 0.509 | |
| Penta_count (%) | | | 0.154 | |
| 1 dose | 112 (19.9) | 109 (18.8) | | |
| 2 doses | 268 (47.6) | 251 (43.3) | | |
| 3 doses | 183 (32.5) | 220 (37.9) | | |
| Measles_Administration (%) | 525 (86.9) | 542 (88.3) | 0.761 | |
| Full Immunization without Birth Doses = Yes (%) | 138 (22.5) | 160 (25.4) | 0.259 | exact |

There were no predictors differentially distributed between treatment and control groups prior to sample enrollment, as per the modified significance testing threshold of 0.000725 (determined through the Bonferroni Correction to adjust for 69 tests). Baseline coverage results show higher proportions of mothers in Control were educated past the 6$^{th}$ grade compared to mothers in Treatment (26.9% vs. 25.2%). A higher proportion of mothers in Treatment were part of the Scheduled Caste category compared to mothers in Control (7.8% vs. 3.8%). A higher proportion of children in Treatment were born outside of facilities compared to children in Control (15.3% vs. 13.2%). There were fewer home visits for postnatal care in the Treatment group when compared to the control group. Health behaviour outcomes were balanced when considering seeking antenatal care checkups, receipt of TT vaccine and IFA, breastfeeding, immunizations (although there were more pentavalent 3 shots in Treatment when compared to Control).

*a) Full Infant Immunization Status using Recall at Endline*

*Table 10. Baseline and Endline Differences in Full Immunization Status from Recall*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| | **Immunized** | **Eligible** | **%(95%CI)** | **Immunized** | **Eligible** | **%(95%CI)** | **95% CI (difference)** |
| **Baseline Coverage Sample**[13] | 160 | 631 | 25.4% | 138 | 612 | 22.5% | -1.93-7.54% |
| **Endline Enrollment Sample** | 837 | 1152 | 72.7% | 629 | 1091 | 57.7% | 11.1-18.9% |
| **Difference (Endline – Baseline)** | | | +47.3% (43.0-51.6%) | | | +35.1% (30.7-39.5%) | |
| **Difference in Difference** | | | +12.2% (3.5-20.9%) | | | | |

**Differences are unadjusted*

The unadjusted mean proportion of infants full immunized per recall was 72.7% in Treatment and 57.7% in Control at Endline, representing a respective increase from baseline of 47.3 percentage points and 35.0 percentage points. The unadjusted difference-in-difference in full infant immunization was 12.2 percentage points higher in Treatment, which was both programmatically meaningful and statistically significant as per the cluster-adjusted, minimum detectable difference threshold at the $p < 0.05$ level. The unadjusted odds ratio for the Treatment group was 2.03 (95% CI 1.60-2.58). Randomization by design serves to balance both measured and unmeasured confounders between the study group, to isolate the true effect of the intervention on the outcome of interest. However, by chance, randomization assignment may lead to differential distributions of secondary exposures which may confound the relationship between study group and outcome. The abbreviated baseline randomization group is provided in full below in the Appendix. Notable differences at baseline between the randomized groups included: *electricity at home, income bracket, land ownership, self-ownership of the mobile phone, roof type, latrine type, and ownership of various fixed assets.* These variables were summarized in the SES Index and used to adjust for the outcome in the final model. The sample followed for the recall outline did face attrition. To check for differential attrition in the outcome-specific sample on any particular

[13] Note that the Baseline Coverage Sample was distinct from the Baseline Enrollment Sample (from which the Endline Enrollment Sample was derived). See page 19 for clarification on the difference in samples.

covariate, a second balance table was made for endline, also including potential differences related to ANMs.

## *Table 11. Endline Differences in Study Groups for Full Immunization Status by Recall*

| | Control | Treatment | p Value[14] |
|---|---|---|---|
| *Study Participants at Endline* | **1091** | **1152** | |
| **ANM Factors** | | | |
| ***Demographics*** | | | |
| ANM Caste | | | 3.0E-12 |
| GEN | 338 (31.0) | 271 (23.5) | |
| SC | 272 (24.9) | 451 (39.1) | |
| ST | 481 (44.1) | 430 (37.3) | |
| ANM Marital Status | | | 2.0E-7 |
| Live-in, Divorced | 4 (0.4) | 8 (0.7) | |
| Married | 1037(95.1) | 1096 (95.1) | |
| Never married | 23 (2.3) | 47 (4.1) | |
| Widowed | 27 (2.5) | 1 (0.1) | |
| ***Phone use*** | | | |
| ANM Access_to_Smartphone_True = 1 (%) | 404 (37.0) | 593 (51.5) | 8.0E-12 |
| Mobile_App_Use: SMS = 1 (%) | 85 (7.8) | 0 (0) | - |
| Mobile_App_Use: WhatsApp = 1 (%) | 251 (23.0) | 448 (38.9) | 7.0E-16 |
| Mobile_App_Use: Youtube = 1 (%) | 837 (76.7) | 693 (60.2) | 6.0E-17 |
| | | | |
| ***Work Processes*** | **Treatment** | **Control** | |
| Hours_per_Camp_Baseline (mean (sd)) | 6.11 (1.12) | 5.71 (1.06) | 2.0E-17 |
| Time_to_Furthest_Camp_1_2hour_Cat = 1 (%) | 132 (12.1) | 257 (22.3) | 2.0E-10 |
| Time_to_Furthest_Camp_Over2hour_Cat = 1 (%) | 50 (4.6) | 18 (1.6) | 5.2E-5 |
| Time_to_PHC_from_House_1_2hour_Cat = 1 (%) | 150 (13.7) | 246 (21.4) | 3.1E-6 |
| Transport_to_Camp_Method_Bike_Cat = 1 (%) | 65 (6.0) | 142 (12.3) | 2.8E-7 |
| Transport_to_Camp_Method_PrivateVehicle_Cat = 1 (%) | 60 (5.5) | 14 (1.2) | 2.7E-8 |
| Connect with ASHA Frequency | | | 2.0E-13 |
| Daily | 186 (17.0) | 109 (9.5) | |
| Weekly | 88 (8.1) | 193 (16.8) | |
| Biweekly | 817 (74.9) | 850 (73.8) | |
| Drop_Out_Tracking_Method: ASHA List = 1 (%) | 738 (67.6) | 673 (58.4) | 7.6E-6 |
| Duties_Outside_Camp: Facility-based care provision = 1 (%) | 677 (62.1) | 602 (52.3) | 3.5E-6 |
| ANM_RCH_Number_Columns (median [IQR]) | 20.00 [13.00, 36.00] | 25.00 [15.00, 40.00] | 8.6E-8 |
| Patient_List_Preparation_Time (median [IQR]) | 30.00 [15.00, 30.00] | 30.00 [25.00, 90.00] | 1.0E-24 |
| Reports that she has complete high-risk list prepared = 1 (%) | 1063 (97.4) | 1066 (92.5) | 2.2E-6 |
| ***Intervention Related Factors*** | | | |
| ANM received at least 1 call from KB= 1 (%) | 0 (0.0) | 482 (41.8) | - |
| Endline Average_Anm_Call_Duration_Cont (median [IQR]) | 0.00 [0.00, 0.00] | 0.00 [0.00, 36.00] | - |
| Prefer_KB_True = 1 (%) | 999 (91.6) | 1142 (99.1) | 2.0E-17 |
| | | | |

[14] "-" corresponds to a p-value less than 1E-10 per the tableone R library

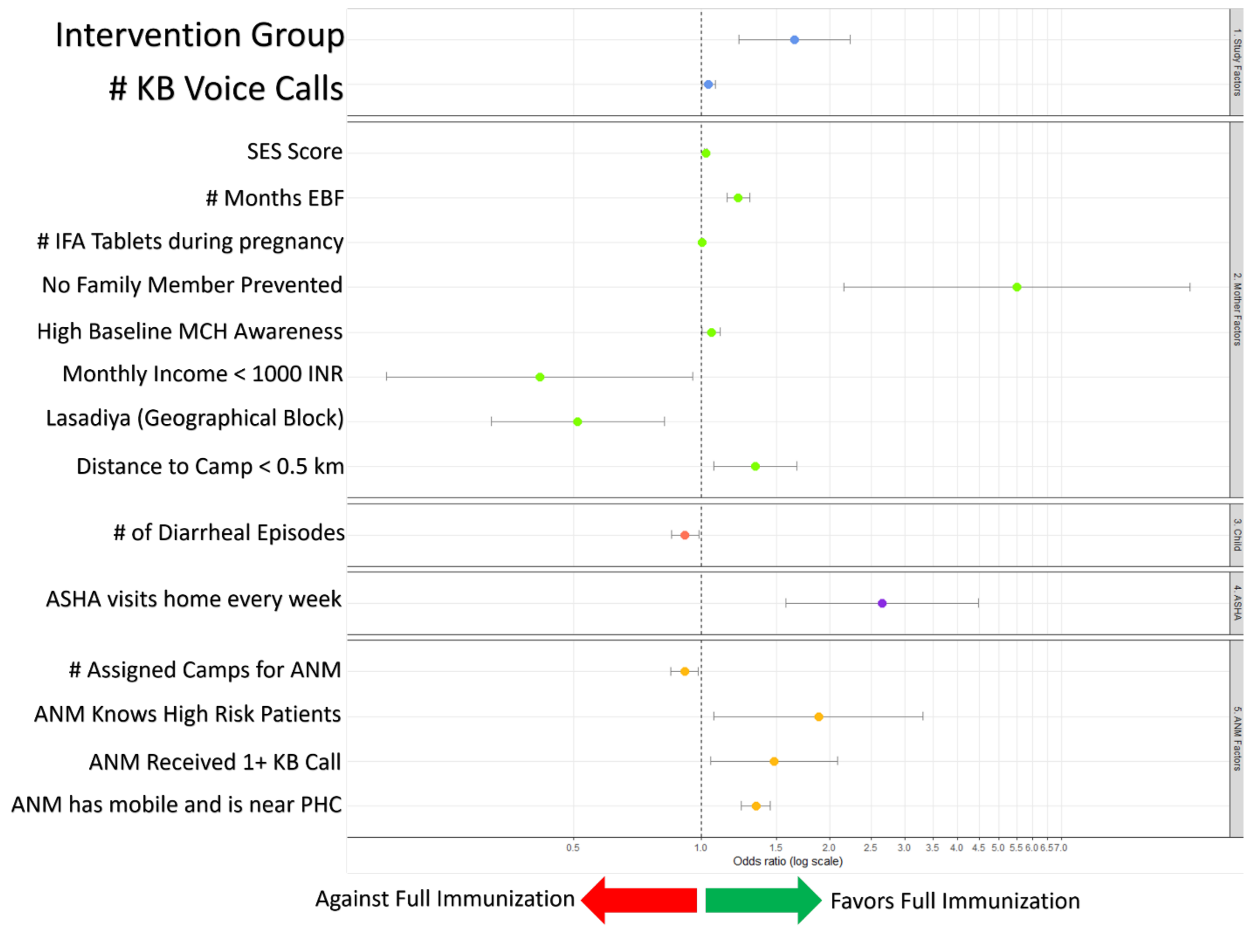
| **Mother Factors** | **Control** | **Treatment** | **p Value** |
|---|---|---|---|
| Monthly Income in INR (%) | | | 4.0E-23 |
| <1000 | 27 (2.5) | 3 (0.3) | |
| 1000-2000 | 98 (9.0) | 20(1.7) | |
| 2001-4000 | 146(13.4) | 105(9.1) | |
| 4001-6000 | 468(42.9) | 489(42.4) | |
| 6001-10000 | 253(23.2) | 400(34.7) | |
| 10000+ | 99(9.1) | 135 (11.7) | |
| Monthly_Saving (median [IQR]) | 750.00 [0.00, 1500.00] | 1500.00 [250.00, 1500.00] | 2.0E-10 |
| Latrine: Jungle/Field = 1 (%) | 983 (90.1) | 965 (83.8) | 1.2E-5 |
| Asha_Visit_After_Birth Distribution (%) | 2.00 [0.00,2.00] | 1.00 [0.00,2.00] | 5.1E-8 |
| ASHA Visit Frequency (%) | | | 1.9E-5 |
| Every day | 7 (0.6) | 5 (0.4) | |
| Every week | 74 (6.8) | 36 (3.1) | |
| Every fifteen days | 180 (16.5) | 137 (11.9) | |
| Every month | 707(64.8) | 823(71.4) | |
| Sometimes | 119 (10.9) | 147 (12.8) | |
| Never | 4 (0.4) | 4 (0.3) | |
| Birth Place (%) | | | 8.4E-8 |
| Home | 143 (13.1) | 163 (14.1) | |
| In transit | 8 (0.7) | 9 (0.8) | |
| Subcenter | 26 (2.4) | 52 (4.5) | |
| Primary Health Center | 387 (35.6) | 317 (27.5) | |
| Community Health Center | 299 (27.4) | 435 (37.8) | |
| Government Hospital | 165 (15.1) | 132 (11.5) | |
| Private Clinic | 61 (5.6) | 43 (3.7) | |
| Delivery_Incentive: Medicines = 1 (%) | 286 (26.2) | 623 (54.1) | - |
| Delivery_Incentive: Supplementary Food = 1 (%) | 254 (23.3) | 593 (51.5) | - |
| Waiting Time at Health Camps (%) | | | 2.0E-6 |
| Less than 5 mins | 115 (10.5) | 51 (4.4) | |
| 5-10 minutes | 344 (31.5) | 363 (31.5) | |
| 10-20 minutes | 256 (23.5) | 280 (24.3) | |
| 20-30 minutes | 157 (14.4) | 189 (16.4) | |
| 30 minutes to 1 hour | 129 (12.2) | 174 (15.1) | |
| More than one hour | 90 (8.2) | 95 (8.2) | |
| TT_Vaccine (mean (sd)) | 1.86 (0.53) | 1.74 (0.47) | 5.0E-9 |
| Child_Age (mean (sd)) | 15.40 (2.10) | 16.29 (1.98) | - |
| ***Intervention Related Factors*** | | | |
| Endline Average_Child_Call_Duration from Backend (median [IQR]) | 0.00 [0.00, 0.00] | 0.00 [0.00, 17.00] | - |
| Endline Total_Child_Calls from Backend (median [IQR]) | 0.00 [0.00, 0.00] | 0.00 [0.00, 2.00] | - |
| Endline Total_Child_Duration from Backend (median [IQR]) | 0.00 [0.00, 0.00] | 0.00 [0.00, 37.25] | - |
| How_Many_Voice_Calls from Recall (median [IQR]) | 0.00 [0.00, 0.00] | 0.00 [0.00, 5.00] | - |
| Mother of child received at least 1 call = 1 (%) | 1 (0.1) | 324 (28.1) | - |
| ***Outcomes*** | **Control** | **Treatment** | |
| PENTA_1 = 1 (%) | 998 (91.5) | 1112 (96.5) | 6.6E-7 |
| PENTA_3 = 1 (%) | 768 (70.4) | 949 (82.4) | - |
| OPV_3 = 1 (%) | 760 (69.7) | 947 (82.2) | - |
| Full_Immunization = 1 (%) | 629 (57.7) | 837 (72.7) | - |

As described previously, a systematic model selection approach was applied to arrive at the final model, which included adjustment for these unbalanced secondary exposures which could potentially confound the relationship between study group and outcome. Base model covariates considered were study group, maternal age, SES score, access to mobile phone, number of antenatal care check-ups, whether the mother received the tetanus vaccine (TT) in pregnancy, maternal awareness score, child awareness, child sex, score, distance to camp, average time spent waiting at the camp, household visit frequency of the ASHA, geographic block, voice calls received by the mother, ANM age, voice calls received by the ANM, and subcenter.

Significant base model covariates, associated with increased full immunization rates in the final model were study group, higher SES score, access to mobile phone, distance to camp less than 0.5km, higher child awareness score, higher maternal awareness score, ASHA visit frequency every week, and higher number of voice calls received by the mother (from recall). Being in the Lasadiya block was associated with a decreased full immunization rate. Other significant final model variable covariates, associated with increase full immunization included: higher number IFA tablets consumed during antenatal care, higher number months spent exclusively breastfeeding, no family member ever prevented immunization, ANM having knowledge of high risk patients, ANM having received at least one KB call, and the ANM having a mobile phone and being near to the PHC. Covariates associated with decreased full immunization included monthly income less than 1000 INR, number of diarrheal episodes of the child. and number of assigned camps to the ANM. Although more Endline surveys were conducted in June and May for Treatment (as opposed to April and May for Control), temporal confounding was not found to be significant (See Appendix for further details).

Without adjustment, those randomized to treatment were 2.03 (95% CI 1.60-2.58) times more likely to be fully immunized. After adjusting for confounders and effects of clustering, mothers randomized to Subcenters which received the Khushi Baby intervention were 1.66 times (95% CI 1.23-2.24) as likely to report full infant immunization at the Endline Survey than those mothers randomized to Control Subcenters. For every additional voice call received from the Khushi Baby system, mothers were 3.8% (95% CI 1.6-7.7) more likely to have their child complete full immunization after adjusting for relevant confounding variables. For every 10-point increase on the SES Index of the mother (range from 5-84), there was an associated 22.6% increase (95% CI 14.6%-30.7%) in odds of full immunization of the child. For every additional 100 tablets consumed during pregnancy, the likelihood of the child completing full immunization increased by 0.26% (95% CI 0.03-0.49). For every additional month spent breastfeeding the likelihood of the child completing immunization increased by 22.1% (95% CI 15.0-29.9). Mothers who were not prevented from having their child immunized by a family member were 5.50 times (95% CI 2.16-14.0) more likely to have the child complete full immunization. For every point increase in MCH awareness score at baseline (principal component score) there was an associated 5.38% (95% CI 0.44-10.6) increase in full immunization rate. Mothers who had a monthly income less than 1000 INR had 0.417 times (95% CI 0.182-953) the odds of completing full immunization for their child when compared to mothers making more than 1000 INR. Mothers living in the Lasadiya block had 0.512 times (95% CI 0.320-0.820) the odds of completing full immunization of their child when compared to mothers from other blocks. Mothers who lived within 0.5 km of the health camp had 1.34 times (95% CI 1.07-1.68) the odds of completing full immunization for their child when compared to mothers living further than 0.5 km from the camp. For every additional diarrheal episode a child experienced, the child was 8.85% (95% CI 2.2-15.1) less likely to complete full immunization. Mothers who had ASHAs visiting their home weekly were 2.66 times (95% CI 1.58-4.48) more likely to have their child full immunized. For every additional camp assigned to the ANM of the mother, the child was 8.8% (95% CI 1.8-15.4) less likely to be fully immunized. Mothers with ANMs that reported knowledge of their highest risk patients were 1.88 times (95% CI 1.07-3.31) more likely to have their child fully immunized. Mothers of ANMs who received at last 1 voice call from the KB system were 1.48 times (95% CI 1.05-2.09) more likely to have their child fully immunized. Mothers with ANMs with mobile phone and near the PHC were 1.34 times (95% CI 1.24-1.45) more likely to have their child fully immunized.

*Figure 7. Significant Covariates associated with Full Immunization (Recall)*




Sensitivity analysis was performed by comparing the results for full immunization by 9 months of age between Treatment and Control, and by running the same analysis using MAMTA card data for both immunization at 12 months and 9 months of age (Outcome 4B), and by repeating within-treatment analysis for data from the Pendant group (Outcome 4C). Results consistently demonstrated that being randomized to Treatment resulted in the mother having a higher odd of her infant completing full immunization when compared to mothers randomized to Control, under an intention-to-treat analysis framework. It's worth noting that there were differences in the point estimates for the full immunization rate between recall and the MAMTA card. The lower immunization rate on MAMTA cards could be explained by multiple factors. For example, ANMs may have forgetten to fill the details of the encounter on both the MAMTA card and in the RCH registers. Oral recall on the other hand may have been influenced by selective memories of the mother, who may be more likely to exaggerate services received to avoid stigma or perceived repercussions for failure to receive all required services.

*Table 12. Full Infant Immunization by Recall and MAMTA card*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| | **Immunized** | **Eligible** | **%** | **Immunized** | **Eligible** | **%** | **Difference (95%CI)** |
| Full Immunization at 12 Months (Recall) | 837 | 1152 | 72.7% | 629 | 1091 | 57.7% | 11.1-18.9% |
| Full Immunization at 12 Months (MAMTA card) | 632 | 921 | 68.6% | 468 | 815 | 57.4% | 6.66-15.7% |
| Full Immunization at 9 Months (Recall) | 849 | 1173 | 72.4% | 649 | 1127 | 57.6% | 10.9-18.7% |
| Full Immunization at 9 Months (MAMTA card) | 643 | 938 | 68.6% | 480 | 847 | 56.7% | 7.41-16.4% |

Subgroup analysis was performed for the following strata: socioeconomic status quartiles (derived from our index, which was verified for normal distribution), baseline maternal health awareness halves, baseline child awareness halves, and distance to camp. Interaction terms were chosen using two necessary criteria: that the individual interaction terms must be independently significantly associated with the outcome and that conceptually the rationale for a synergistic effect must be justified. In this case, it was hypothesized that mothers in the highest quartile may receive higher benefits from the intervention than those mothers in lowest quartile due to the fact that mothers in the highest quartile may be more educated and willing to learn and more likely to have mobile phone to receive calls which may be charged due to electricity present at home. Similarly, it was hypothesized that mother with a higher baseline MCH awareness score would be more able to learn and respond to community engagement efforts and that mothers closer to the camps would be more likely to attend the camp in the case that they received a reminder.

The table below demonstrates the strata specific proportions of outcomes between Treatment and Control groups with the odds ratio, confidence interval, and p-value of the interaction term when added to the model. None of the interaction terms reached significance. Mothers with higher maternal health awareness in Treatment were 1.48 times more likely to have their child immunized compared to mothers in Control with low baseline maternal health awareness (p=0.058). As such there was insufficient evidence to suggest that the effect of the treatment was partial to any of the mentioned subgroups below.

*Table 13. Subgroup Analysis for Full Immunization Status by Recall*

| Sub Group | Treatment | | | Control | | | | |
|---|---|---|---|---|---|---|---|---|
| ***x = Immunized*** ***X = Eligible*** | **x** | **X** | **%** | **x** | **X** | **%** | **Odds Ratio (95% CI)** | **P Value** |
| SES Quartile 1 | 175 | 273 | 64.1% | 151 | 304 | 49.7% | * | * |
| SES Quartile 2 | 215 | 307 | 70.0% | 173 | 307 | 56.4% | 0.988 (0.587-1.66) | 0.964 |
| SES Quartile 3 | 208 | 272 | 76.5% | 124 | 223 | 55.6% | 1.41 (0.801-2.48) | 0.234 |
| SES Quartile 4 | 239 | 300 | 79.7% | 181 | 257 | 70.4% | 0.814 (0.458-1.45) | 0.486 |
| | | | | | | | | |
| Maternal Awareness Bottom 50% | 435 | 625 | 69.6% | 359 | 612 | 58.7% | * | * |
| Maternal Awareness Top 50% | 402 | 527 | 76.3% | 270 | 479 | 56.4% | 1.48 (0.987-2.21) | 0.058 |
| Child Awareness Bottom 50% | 411 | 584 | 70.4% | 304 | 570 | 53.3% | * | * |
| Child Awareness Top 50% | 426 | 568 | 75.0% | 325 | 521 | 62.4% | 0.741 (0.493-1.11) | 0.149 |
| Distance to Camp >= 0.5km | 575 | 807 | 71.3% | 423 | 765 | 55.3% | 0.849 (0.548-1.32) | 0.465 |
| Distance to Camp <0.5km | 262 | 345 | 75.9% | 206 | 326 | 63.2% | * | * |

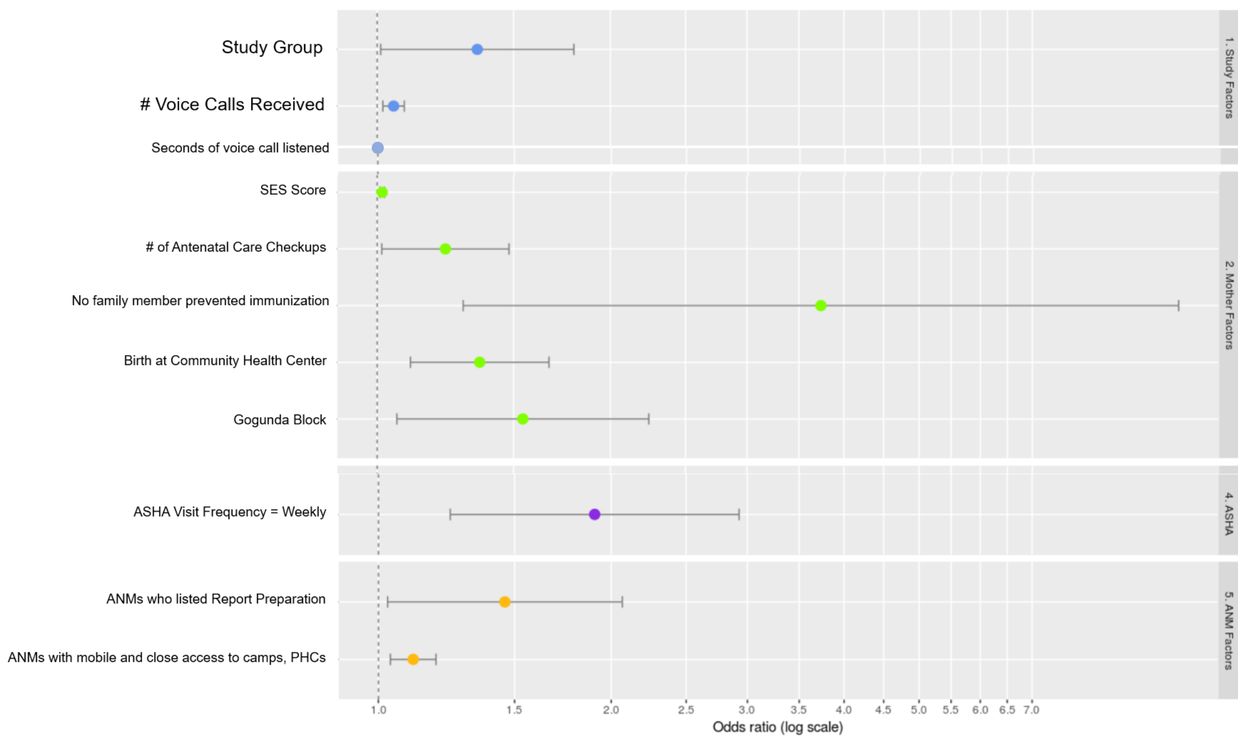
*b) Full Infant Immunization Status by MAMTA card at Endline*

The purpose of this analysis was to corroborate the outcome estimates and conclusions of study group differences with another common approach to measuring full immunization – use of the government-issued MAMTA card. Mothers who received and retained their MAMTA card until Endline were eligible for this analysis. Randomization balance was checked for the Endline sample. Notable differences can be seen in Appendix Table 2. A systematic model selection approach was applied to arrive at the final model, which included adjustment for these unbalanced secondary exposures which could potentially confound the relationship between study group and outcome. The same base model variables included in the Outcome 4A were considered here.

Without adjustment for confounders, mothers in the Treatment group were 1.70 (95% CI 1.27-2.28) times more likely to have their child fully immunized. After adjusting for confounders and effects of clustering, mothers randomized to Subcenters which received the Khushi Baby intervention were 1.35 times (95% CI 1.10-1.67) as likely to report full infant immunization at the Endline Survey than those mothers randomized to Control Subcenters. For every additional voice call received from the Khushi Baby system, mothers were 5.0% (95% CI 1.8-8.4) more likely to have their child complete full immunization after adjusting for relevant confounding variables. For every additional minute the mother spent listening to a KB-generated voice call (as verified by the Twilio database), the child was an 8.27% increase (95% CI 3.12-13.4) in the likelihood of the child be fully immunized by Endline.

For every 10-point increase on the SES Index of the mother (range from 5-84), there was an associated 15.5% increase (95% CI 8.8-22.2) in odds of full immunization of the child. Mothers living in the Gogunda block had 1.54 times (95% CI 1.06-2.24) the odds of completing full immunization of their child when compared to mothers from other blocks. Mothers who had ASHAs visiting their home weekly were 1.90 times (95% CI 1.24-2.93) more likely to have their child full immunized. For every additional antenatal care checkup completed by the Mother, the child was 22.5% (95% CI 1.44-48.0) more likely to be immunized. If the child was born at a Community Health Center, the child was 1.35 times (95% CI 1.10-1.67) more likely to be fully immunized compared to children born elsewhere (home, transit, PHC, or hospital). Mothers who were not prevented from having their child immunized by a family member were 3.74 times (95% CI 1.29-10.8) more likely to have the child complete full immunization. Mothers who had ANMs who listed Report Preparation as a duty outside of camp were 1.46 times (95% CI 1.03-2.07) more likely to have their child with a full immunization status as per the MAMTA card. Mothers who had ANMs who were closer to camps, PHCs, and had access to a mobile phone (principal component group score) were 1.09 times (95% CI 1.04-1.19) more likely to have their child immunized.

*Figure 8. Adjusted Odds covariates associated with Full Immunization (MAMTA)*




Subgroup analysis was intended for the same strata as mentioned in Outcome 4A. However, maternal awareness, child awareness, and distance form camp were not significantly associated with the outcome as independent strata, so only the interaction between SES quartile and study group was considered. The interaction term was again not found to be significant and resembled findings from Outcome 4A.

*Table 14. Subgroup Analysis for Full Immunization by MAMTA card*

| | Treatment | | | Control | | | | |
|---|---|---|---|---|---|---|---|---|
| ***x = Immunized X = Eligible*** | x | X | % | x | X | % | Odds Ratio (95% CI) | P Value |
| SES Quartile 1 | 133 | 273 | 48.7% | 105 | 304 | 34.5% | | |
| SES Quartile 2 | 161 | 307 | 52.4% | 123 | 307 | 40.1% | 0.798 (0.481-1.32) | 0.388 |
| SES Quartile 3 | 148 | 272 | 54.4% | 100 | 223 | 44.8% | 0.749 (0.436-1.28) | 0.292 |
| SES Quartile 4 | 190 | 300 | 63.3% | 140 | 257 | 54.5% | 0.692 (0.404-1.19) | 0.180 |

c) Full Infant Immunization Status by Khushi Baby Pendant at Endline

The purpose of this analysis was to corroborate Outcomes 4A and 4B with treatment-specific data. As seen in Outcome 2A and 2B, the data stored on the pendant was more likely to be complete and consistent with the Khushi Baby backend. By design, vaccination statuses also contained the date of administration automatically, which was not the case with the MAMTA card. Compared to recall (72.7% and 72.4% for 12 months and 9 months respectively) and MAMTA card (68.6% for both 12 and 9 months), full immunization coverage was lower roughly by 5-10% on the pendant. When just looking at the 778 respondents with pendants, 601/778 (77.2%) reported full immunization by recall and 479/778 (61.5%) were fully immunized as per the MAMTA card at 12 months.

### *Table 15. Full Immunization Status by Khushi Baby Pendant*

| | Immunized | Eligible | % |
|---|---|---|---|
| Full Immunization by 12 months (Pendant) | 481 | 754 | 63.8% |
| Full Immunization by 9 months (Pendant) | 488 | 767 | 63.6% |

The sample considered for regression analysis were mothers who received pendants and were able to present the verified pendant at Endline. Because this was a treatment-specific analysis, randomization balance was not needed. A systematic model selection approach was applied to arrive at the final model.

After multivariate regression, for each additional voice calls received by the mother as per the KB backend, there was a 1.06 (95% CI 1.01-1.11) times higher likelihood of completing full immunization. A ten-point increase in SES score was associated with a 14.7% (95% CI 0.089-28.9%) in the odds of full immunization. For every additional pneumonia episode experienced by the child there was a 12.1% (95% CI 0.08-21.2%) decrease in the likelihood of full infant immunization. For every grade increase in the ANM's education there was a 1.33 (95% CI 1.17-1.50) times higher likelihood of the outcome. Mothers who had ANMs who faced difficulty arranging the register were 40.4%(95% CI 9.7-60.7%) less likely to complete immunization. Mothers with ANMs who tracked dropout lists with the RCH register were 45.2% (95% CI 14.1-65.0%) less likely to complete full immunization.

When considering the differences between immunization rates across all three methods, figures 10-12 below demonstrate the higher rates of immunization in Treatment (Figure 11) compared to Control (Figure 10) across individual rates of vaccines. Highest drop-outs take place between

PENTA 3 and Measles for both Control and Treatment using MAMTA card data, but between PENTA 2 and 3 using KB Pendant and KB Backend Data (Figure 12).

*Figure 9. Immunization Coverage Rates by MAMTA Card, Recall in Control Subcenters*

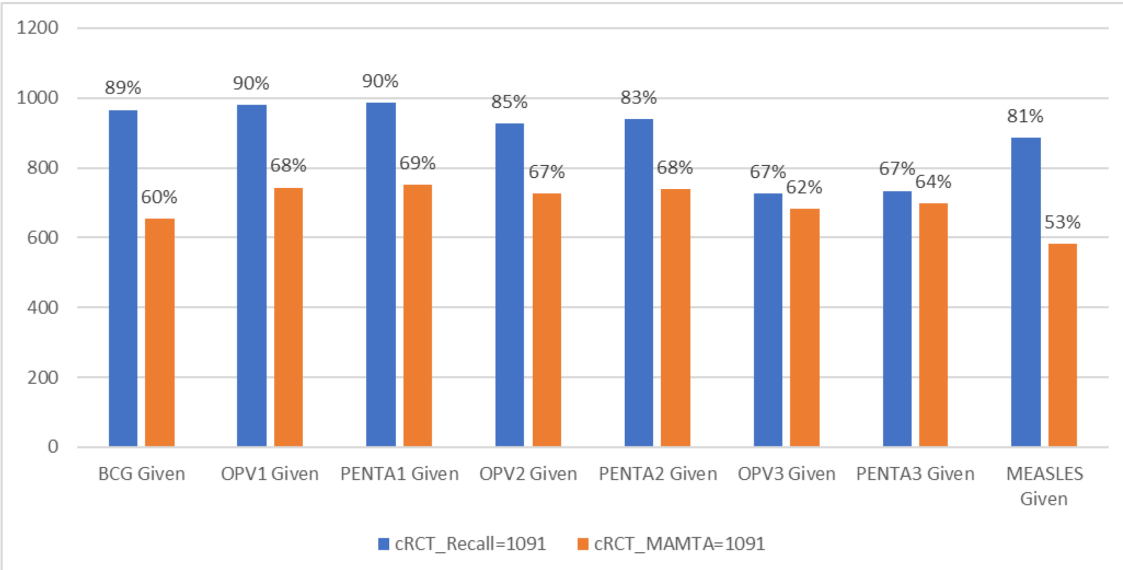



*Figure 10. Immunization Coverage Rates by MAMTA Card, Recall, and Pendant in Treatment Subcenters*

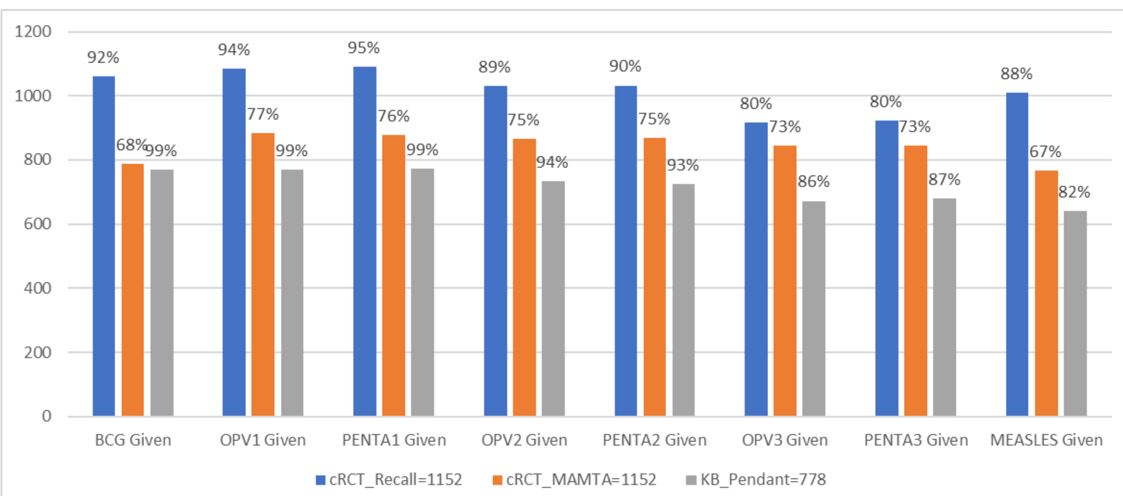
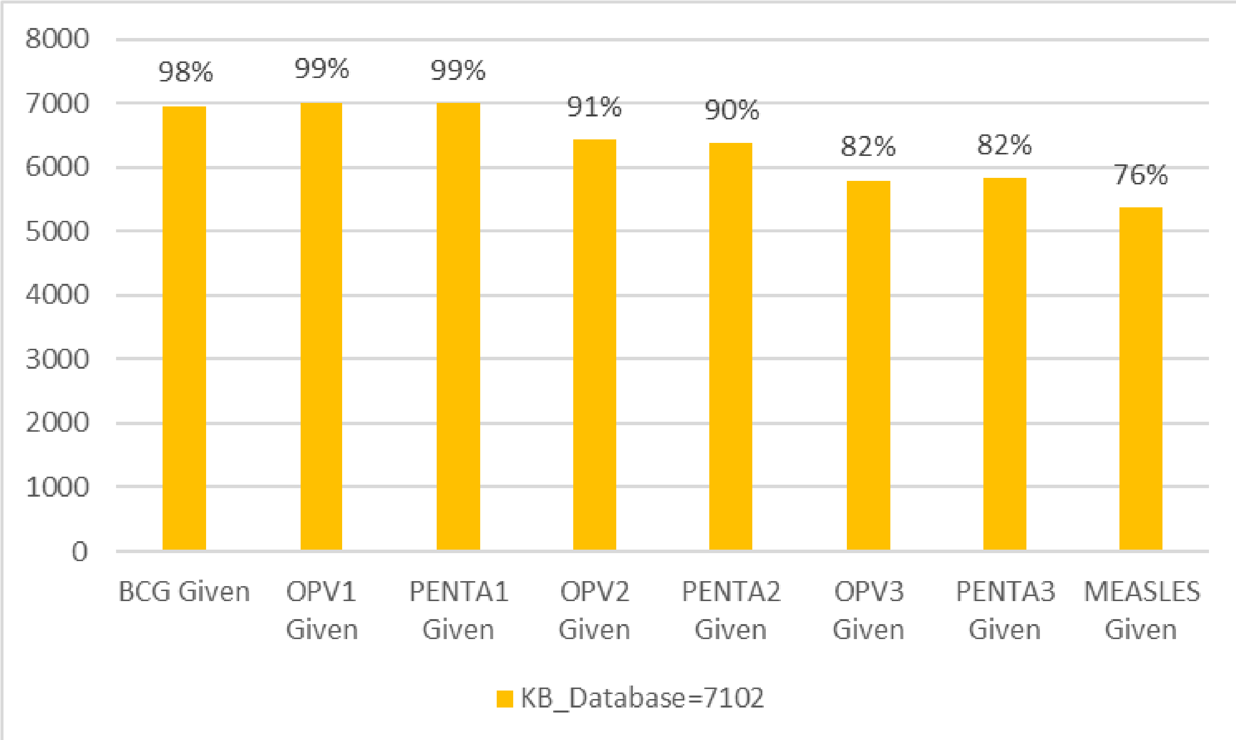



*Figure 11. Immunization Coverage Rates by Khushi Baby Backend from February 2017 to October 2018*




*d) PENTA 1-3 Coverage by Recall*

*Table 16. PENTA1-3 Completion Status by Recall at Endline*

| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| | **Immunized** | **Eligible** | **% (95% CI)** | **Immunized** | **Eligible** | **% (95%CI)** | **Difference (95%CI)** |
| **Baseline** | | | | | | | |
| PENTA 1-3 Coverage (Recall) | 220 | 631 | 34.9% | 183 | 612 | 29.9% | -0.23-10.2% |
| **Endline** | | | | | | | |
| PENTA 1-3 Coverage (Recall) | 910 | 1152 | 79.0% | 716 | 1091 | 65.5% | 9.7-17.0% |
| PENTA 1,2 | 1032 | 1152 | 89.6% | 940 | 1091 | 86.1% | 0.72-6.1% |
| PENTA 1 | 1090 | 1152 | 94.6% | 986 | 1091 | 90.3% | 2.1-6.4% |
| PENTA1 to 3 Dropout | 180 | 1152 | 15.6% | 270 | 1091 | 24.7% | -12.4 - -5.8% |
| **Difference (Endline-Baseline)** | | | +48.5% (39.7-48.5%) | | | +39.7% (31.1-40.3%) | |
| **Difference-in-Difference** | | | +8.4% (-0.59-17.4%) | | | | |

**Differences are unadjusted*

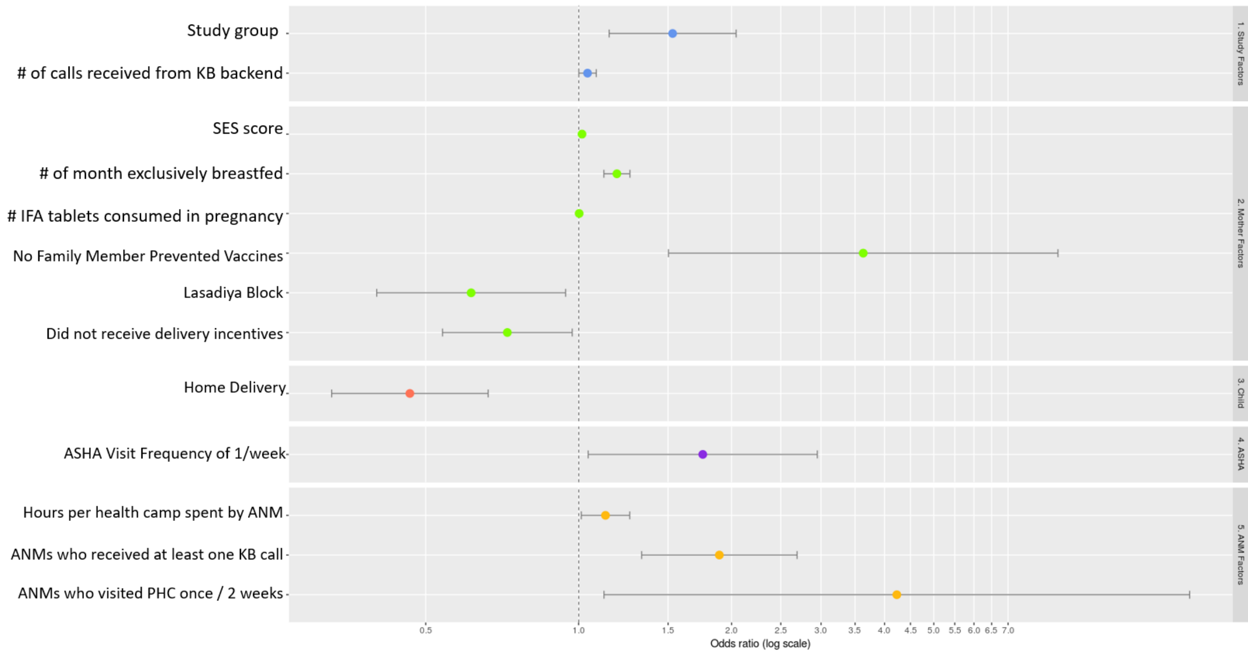
The purpose of this analysis was to investigate whether the results for full immunization were consistent when considering pentavalent series coverage. The pentavalent series is typically due for completion by 180 days after birth and serves as a prerequisite to achieving the full immunization outcome. Like the results for full infant immunization, summary statitstics demonstrated that the Treatment group performed significantly higher than the Control group with respect to Penta 1-3 completion (14.5 percentage point increase) and Penta 1-3 Dropout (9.1

percentage point decrease). Results showed an increase in Penta 1-3 completion for the Treatment group from an unadjusted difference-in-difference comparison, which factored in the change from baseline in each study group (8.4 percentage point). These results were comparable to the findings for full immunization (12.2 percentage point difference-in-difference).

Several variables were differentialy distributed between the study groups such as child age (15.99 months in Treatment vs. 15.04 months in control), birth place (greater proportion in Community Health Center for Treatment vs. Primary Health Center in Control), ANM Caste composition, ANM access to network (greater proportion in Treatment), Monthly income between 1000-2000 INR (greater proportion in Control), and waiting time at the health camp les than 5 minutes (greater proportion in Control). Covariates independently associated with the pentavalent coverage outcome included: being visited midline, not being prevented by a family member, caste category, being exclusively breastfed, higher number of IFA tablets consumed, higher education score, higher husband education score, higher transportation score, and higher SES score.

The unadjusted odds of mothers in Treatment completing full pentavalent coverage was 2.04 (95% CI 1.57-2.64) times that of mothers in Control completing full pentavalent coverage. After adjusting for confounding in the multivariate regression model, mothers randomized to Treatment subcenters were 1.53 (95% CI 1.15-2.04) times more likely to have completed the pentavalent series when compared to mothers in Control subcenters. For every additional voice call received by the mother (confirmed by the Twilio backend), the odds of pentavalent series completion increased by 4.12% (95% CI 0.07-8.34%). For every 10 point increase in SES Score of the mother there was a 15.4% (95% CI 7.64-23.4%) increase in the odds of pentavalent immunization. Mothers who exclusively breastfed were 1.19 (95% CI 1.12-1.26) times more likely to complete the pentavalent series for their child. For every additional 100 IFA tablets consumed by the mother during pregnancy, the odds of pentavalent vaccination increased by 2.51% (95% CI 0.22-4.80%). Mothers who did not have a family member prevent vaccination were 3.63 (95% CI 1.50-8.78) times as likely to complete the pentavalent series for their child. Mothers who did not receive incentives during the time of delivery were 0.724 (95% CI 0.540-0.971) times as likely to have completed the pentavalent series. Mothers from Lasadiya were 0.614 (95% CI 0.401-0.942) times as likely to have completed the pentavalent series. Mothers who had home births were 0.465 (95% CI 0.327-0.664). Mothers with ASHAs who visited once per week were 1.75 (95% CI 1.04-2.95) times more likely to complete the pentavalent series. For every additional hour spent for camp activities by the ANM, the mother was 12.9% (95% CI 1.21-26.1%) higher. Mothers who were randomized to ANMs who visited the PHC on a fortnightly basis were 4.23 (95% CI 1.12-15.9) than those who had ANMs visiting on a different cadence. Mothers randomized to ANMs who received at least one voice call from the KB backend system (confirmed via the Twilio database as received by the user), were 1.89 (95% CI 1.33-2.69) times more likely to complete the pentavalent series compared to their counterparts. Subgroup analyses for SES quartile, maternal and child awareness scores, camp distance, and being part of the baseline ANM linelist were all found to be not significant for effect modification.

*Figure 12. Adjusted odds ratios for covariates significantly associated with Pentavalent 1-3 Completion by Recall*




## Khushi Baby System Effect on Health Outcomes

*Table 17. Health Outcome Differences between Treatment and Control Groups*

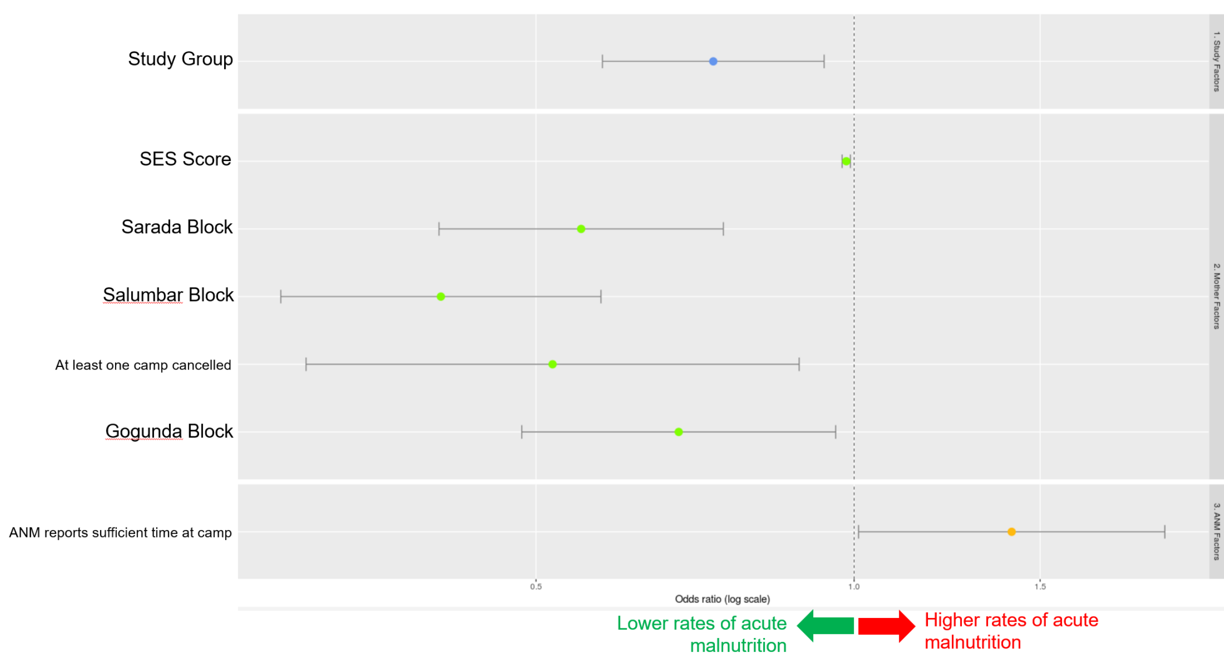
| | **Treatment** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|
| | **Immunized** | **Eligible** | **%** | **Immunized** | **Eligible** | **%** | **p-value** |
| Infant Mortality Rate | 23 | 1243 | 18.50 | 21 | 1192 | 17.62 | 0.870 |
| Moderate Acute Malnutrition (MUAC 11.5-12.5 cm) | 156 | 1122 | 13.9% | 187 | 1045 | 17.9% | 0.011* |
| Severe Acute Malnutrition (MUAC <11.5 cm) | 54 | 1122 | 4.8% | 57 | 1045 | 5.5% | 0.498 |
| Hospital Admission of the Infant | 1009 | 1189 | 84.86% | 971 | 1148 | 84.58% | 0.851 |
| Hospital Admission due to Diarrhea | 786 | 1243 | 63.2% | 753 | 1187 | 63.4% | 0.917 |
| Hospital Admission due to suspected pneumonia | 479 | 1243 | 38.5% | 448 | 1187 | 37.7% | 0.687 |

Both endline Treatment and Control groups reported comparable rates in infant hospitalization (84.86%, 84.58%), hospitalization diarrhea (63.2%, 63.4%), and hospitalization pneumonia (38.5%, 37.7%), respectively. There was a statistically significant difference between groups for Moderate Acute Malnutrition with children in Treatment subcentres having a MAM rate of 13.9% compared to children in Control subcenters having a MAM rate of 17.9% (p = 0.011). The unadjusted odds for the Treatment group on the MAM outcome was 0.766 (95% CI 0.582-1.009), which approached significance.

To check for confounders between the study group and MAM outcome, a randomization balance table was made. Notable differences in differential distribution of secondary exposures can be found in Appendix Table 3. The same systematic approach for outcomes 4A and 4B was applied to narrow the list of covariates included in the final regression model, used to determine the pure fixed effect of the study group.

Without adjustment for confounders, mothers randomized to treatment were 0.766 times (0.582-1.01) less likely to have their infant have Moderate Acute Malnutriton. After adjusting for confounders and effects of clustering, mothers randomized to Subcenters which received the Khushi Baby intervention were 0.736 times (95% CI 0.577-0.937) as likely to have children who had moderate acute malnutrition at endline. Mothers living in Sarada were 0.551 times (95% CI 0.404-0.752) as likely to have their child have MAM. Mothers living in Salumbar were 0.406 times (95% CI 0.286-0.576) as likely to have their child have MAM. Mothers in Gogunda were 0.682 times (95% CI 0.484-0.961) as likely to have their child have MAM. For every one-point increase in SES score of the mother, there was an associated 1.7% (95% CI 0.8-2.6) decrease in the likelihood of the child having MAM. Surprisingly, mothers who faced at least one cancelled camp were 0.518 times (95% CI 0.303-0.887) as likely to have their child have MAM. Also it was unexpected to observe that mothers with ANMs who reported that they had sufficient time at their health camp to perform required activities were 1.41 times (95% CI 1.01-1.97) more likely to have their child have MAM.

*Figure 13. Adjusted odds ratios for covariates significantly associated with Moderate Acute Malnutrition*



Subgroup analysis was intended for the same strata as mentioned in Outcome 4A and 4B. However, maternal awareness, child awareness, and distance form camp were not significantly associated with the outcome as independent strata, so only the interaction between SES quartile and study group was considered. The interaction term was again not found to be significant and resembled findings from Outcomes 4A and 4B.

*Table 18. Subgroup Analysis for Moderate Acute Malnutriton*

| | **Treatment** | | | **Control** | | | | |
|---|---|---|---|---|---|---|---|---|
| ***x = Immunized X = Eligible*** | x | X | % | x | X | % | Odds Ratio (95% CI) | P Value |
| SES Quartile 1 | 43 | 261 | 16% | 60 | 281 | 21% | * | * |
| SES Quartile 2 | 53 | 298 | 18% | 50 | 290 | 17% | 1.57 (0.852-2.92) | 0.146 |
| SES Quartile 3 | 33 | 272 | 12% | 49 | 224 | 22% | 0.725 (0.375-1.40) | 0.337 |
| SES Quartile 4 | 27 | 291 | 9.3% | 27 | 249 | 11% | 1.46 (0.705-3.00) | 0.309 |

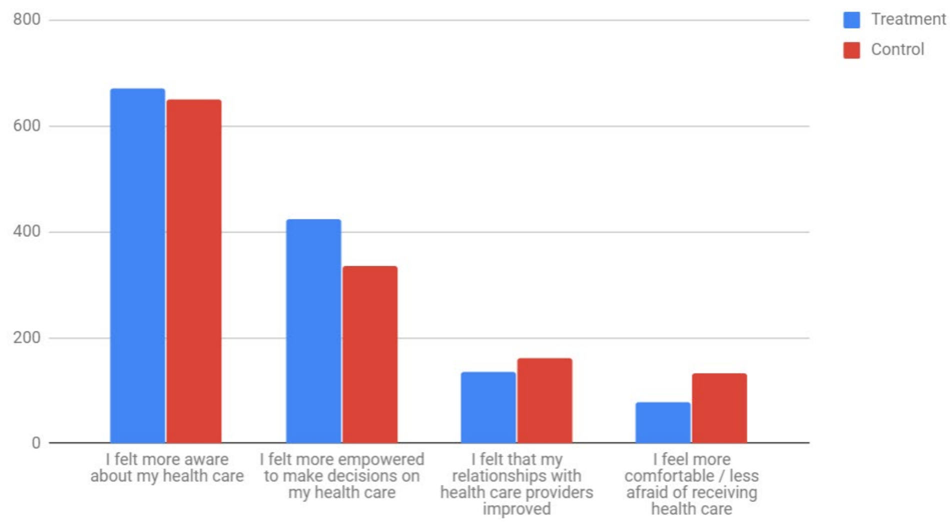
In summary, the significant effects of assignment to the Khushi Baby arm (i.e. study group) on the health behavior and health outcomes can be found in Table 20 below. We can observe for all outcomes, the adjusted odds ratio estimate was less in magnitude when compared to the unadjusted odds ratio estimate, suggesting that other confounders within the multivariate model helped explain the association. The exception to this pattern would be the moderate acute malnutrition outcome, where after adjusting for covariates, the effect size of the odds ratio is increased, and becomes statistically significant, suggesting that the control group had confounders that were otherwise protective against the outcome of moderate acute malnutrition. Many covariates in the multivariate models were shared across outcomes, further corroborating their significance. The number of voice calls received, SES-score, geographical block, ASHA visit frequency, higher MCH awareness (higher IFA consumption, delivery at hospitals, and exclusive breastfeeding), ANM use of mobile phones, and ANM proximity to PHCs were all found to be significant factors in multiple outcome models.

## *Table 19. Summary of Effect Sizes of Study Group on Outcome (Intent to Treat)*

| | **Unadjusted Odds Ratio of Study Group** | | | |
|---|---|---|---|---|
| **Outcome (measurement tool)** | Effect Size | 95% CI | Standard Error | p Value |
| *Full immunization (Recall)* | 1.95 | 1.66-2.33 | 0.09 | 1.18E-13 |
| *Full Immunization (MAMTA)* | 1.28 | 1.37-1.912 | 0.0851 | 1.59E-08 |
| *PENTA 1-3 (Recall)* | 1.95 | 1.616-2.343 | 0.095 | 2.08E-12 |
| *Moderate Acute Malnutrition (MUAC)* | 0.745 | 0.591-0.94 | 0.118 | 0.013 |
| | **Unadjusted Odds Ratio of Study Group, accounting for clustering** | | | |
| *Full immunization (Recall)* | 2.03 | 1.60-2.58 | 0.122 | 6.70E-09 |
| *Full Immunization (MAMTA)* | 1.70 | 1.27-2.28 | 0.149 | 0.00033 |
| *PENTA 1-3 (Recall)* | 2.04 | 1.57-2.64 | 0.132 | 6.67E-08 |
| *Moderate Acute Malnutrition (MUAC)* | 0.766 | 0.582-1.01 | 0.14 | 0.06 |
| | **Adjusted Odds Ratio of Study Group, accounting for clustering and multiple covariates** | | | |
| *Full Immunization (Recall)* | 1.66 | 1.23-2.24 | 0.153 | 0.000998 |
| *Full Immunization (MAMTA)* | 1.35 | 1.011-1.794 | 0.146 | 0.0420 |
| *PENTA 1-3 (Recall)* | 1.53 | 1.15-2.04 | 0.147 | 0.00370 |
| *Moderate Acute Malnutrition (MUAC)* | 0.736 | 0.577-0.937 | 0.123 | 0.0130 |
| **Covariates included in the multivariate models:** | | | | |
| *Full immunization (Recall)* | study group, # of voice calls (self report), SES-score, #IFA consumed, #months breastfeeding, whether family members prevented immunization, maternal child awareness index, monthly income less than 1000 INR, geographical block, frequency of ASHA visits, ANM self-report of having high-risk patient list, ANM receipt of voice reminders from the KB system, ANM possession of mobile phone and distance to Primary Health Centers | | | |
| *Full Immunization (MAMTA)* | study group, # of voice calls (KB Backend), # of minutes listened to KB voice calls (KB Backend), SES-score, # of ANC checkups completed per self report, whether family members prevented immunization, delivery location, geographical block, frequency of ASHA visits, ANM who listed report preparation as a duty outside of camp, ANM possession of mobile phone and distance to camps and PHCs | | | |
| *PENTA 1-3 (Recall)* | study group, # of voice calls (KB Backend), SES-score, #IFA consumed, #months breastfeeding, whether family members prevented immunization, geographical block, receipt of incentives during delivery-time, delivery location, frequency of ASHA visits, ANM reported hours spent at health camp, ANM receipt of voice reminders from the KB system, ANM frequency of visiting PHCs | | | |
| *Moderate Acute Malnutrition (MUAC)* | study group, # of voice calls (self report), SES-score, geographical block, experience of a camp cancellation, ANMs who reported spending sufficent time at the camps | | | |

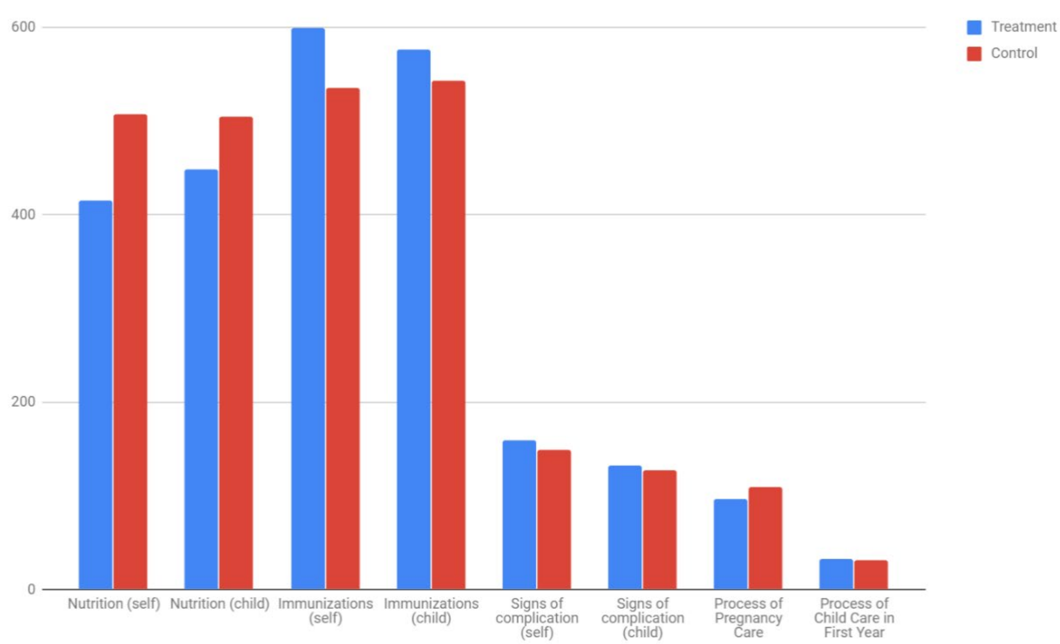
## Khushi Baby System Effect on Attitudes and Perceptions of Beneficiaries

*Figure 14. Attitudes and Perceptions of Mothers with positive experiences (infancy)*



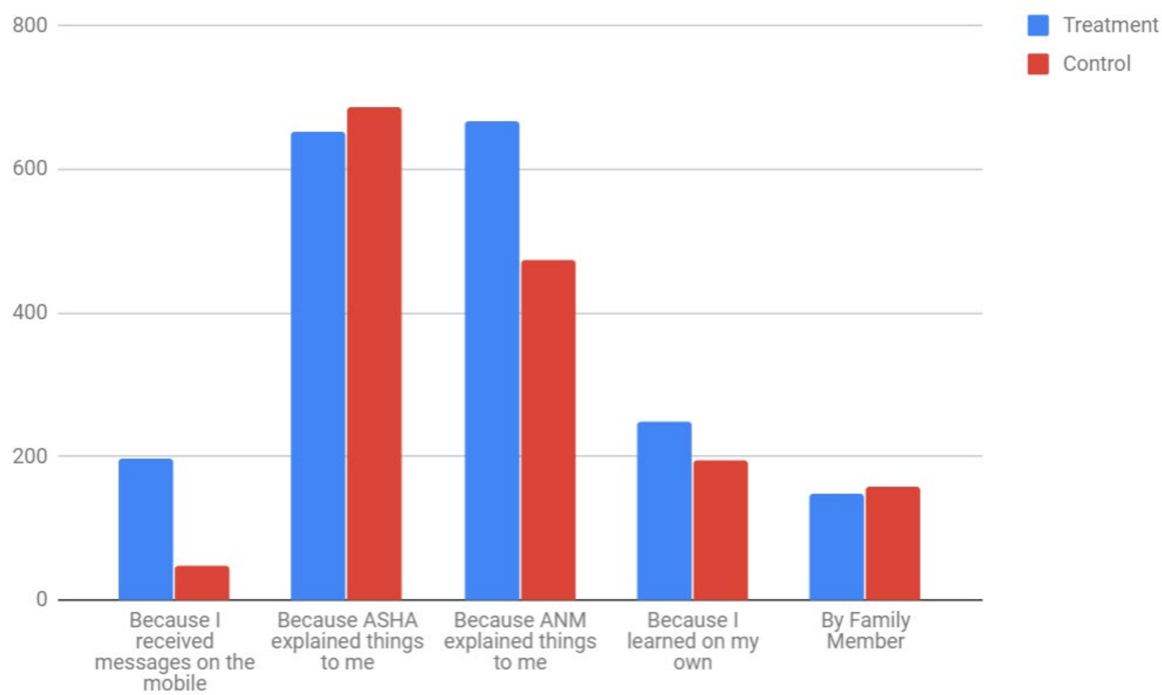

Distributions for the types of positive experience experienced by mothers differed significantly between study groups ($p < 0.00001$, $X^2 = 27.2$). Proportionally more mothers in Treatment felt empowered to make decisions on their health ($p = 0.00052$), whereas proportional more mothers in Control reported feeling more comfort with receiving health care ($p < 0.00001$). There were no significant differences in the underlying distributions of the ways for feeling empowered or the people which health relationships improved between study groups.

*Figure 15. Topics for which mothers reported positive awareness change at endline*



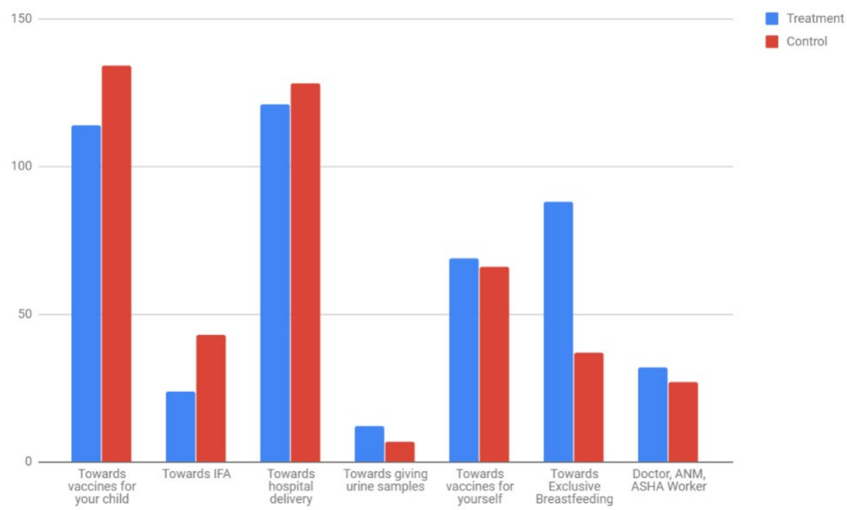
Distributions significantly differed for the types of awareness change ($p = 0.0126$). Mothers in Treatment were more likely to report higher awareness for maternal immunizations than Control ($p = 0.01208$). Mothers in Control were more likely to report higher awareness for maternal nutrition ($p = 0.00228$).

*Figure 16. Reasons driving positive awareness change reported by mothers at endline*



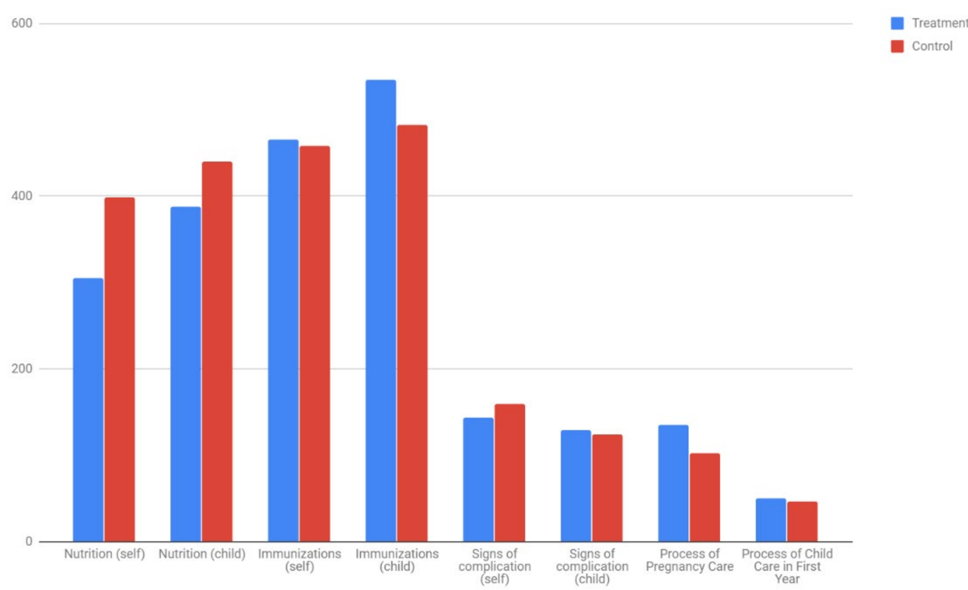

Distributions significantly differed for the drivers of awareness change between study groups ($p<0.00001$). Mothers in Treatment reported in higher proportions that their awareness was increased by mobile messages ($p < 0.00001$), by the ANM interaction ($p = 0.00438$). Mothers in Control reported in higher proportions that their awareness was increased through conversations with the ASHA ($p < 0.00001$).

*Figure 17. Areas in which mothers felt more comfortable towards MCH at endline*



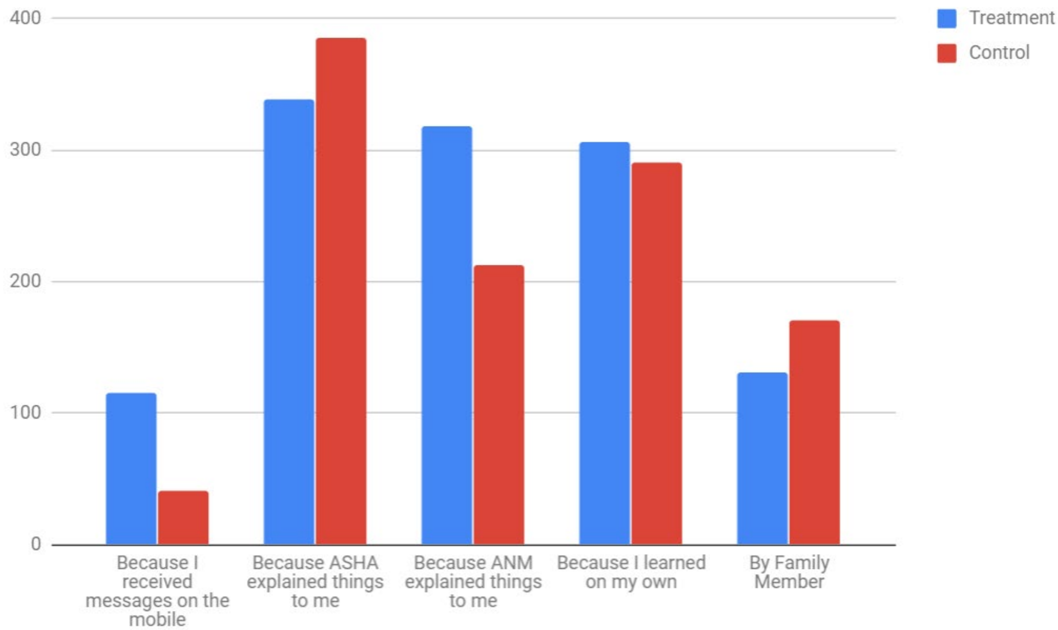

There were significantly distributions in the way mothers felt more comfort between the study groups ($p<0.00005$). Mothers in Treatment reported in significantly higher proportions that they felt more comfortable with exclusive breastfeeding ($p<0.00001$).

*Figure 18. Topics for which mothers reported that the husband's awareness improved*




There were significantly different distributions in the topics for which husband/father's awareness was increased between Treatment and Control ($p = 0.00014$). Mothers in Control reported in higher proportions that their husbands had higher awareness in maternal nutrition ($p = 0.00052$). Mothers in Treatment reported higher rates of husband awareness about child immunizations, but this only approached significance ($p = 0.0164$).

*Figure 19. Reasons driving awareness change in husbands/fathers*




Distributions significantly differed for the drivers of awareness change between study groups ($p<0.00001$). Mothers in Treatment reported in higher proportions that their husband's awareness was increased by mobile messages ($p < 0.00001$), by the ANM interaction ($p = 0.0006$) Mothers in Control reported in higher proportions that their husband's awareness was increased through conversations with the ASHA ($p = 0.00026$) and through family members ($p = 0.00084$).

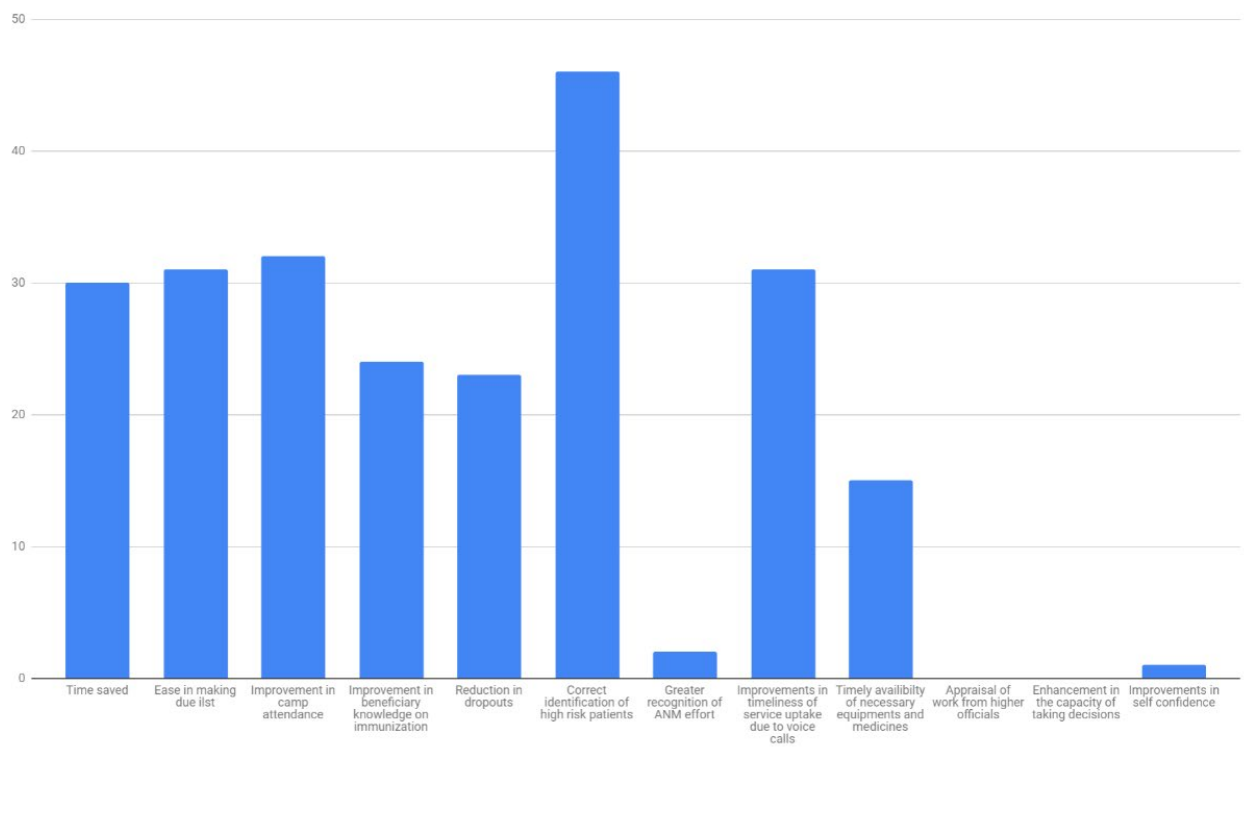
Focus group discussions with mothers generated feedback around three thematic areas: voice calls, the Khushi Baby pendant and biometric, and future ideas and suggestions. While some mothers appreciated the voice call messages, others had perceptions that the call frequency was inconsistent or easily mistaken for an advertisement. Some mothers found the pendant to be convenient and customizable while others felt that the reason for why the pendant needed to be worn was not properly explained. Others found that the whole KB process as a result of the pendant scan may have increased the time spent at the camp. Ideas for improvement included guidance on foods to eat during pregnancy, ability to customize the pendant, and access to more educational content with local language videos. Representative quotes can be found in the Appendix.

*Khushi Baby System Effect on Attitude, Perceptions, and Practices of Health Workers*

At baseline, the highest frequency of self-reported duties included facility based-care, report preparation, coordination with the ASHA, and special government programs. At endline for Treatment ANMs, the highest frequency of self-reported duties included report preparation, home-based visits for postnatal care, attending trainings and meetings, and performing sterilization. Self-reported role at baseline differed from those at endline with more ANMs at endline describing their role as a coordinator (46.4% in endline vs. 35.6% in baseline) with health providers and less as an educator (7.3% vs. 36.8%) and social worker (73.9% vs. 82.8%).

With respect to confidence in carrying out camp procedures, ANMs in Treatment and Control had distributions that were not significantly different for the ability to communicate with mother, ability to record data into the RCH register, decision-making for medication administration, or decision-making for camp-based tests or decision-making ability for vaccine administration. There was no statistically significant difference between time spent per camp in Treatment (5.48 hours) and Control ANMs (5.45 hours).

*Figure 20. Benefits from KB system reported by ANMs*




When presented an open-ended question without suggested options regarding the system's positive aspects, 46/69 ANMs reported that the KB System helped them correctly identify high risk patients. 32/69 ANMs attributed the KB system towards an improvement in camp attendance by beneficiaries. 31/69 ANMs reported that the KB App assisted them with making their due list.

62/69 (90%) of ANMs who used the app reported that they would prefer to continue to use the system moving forward. 68/77 (88%) of ANMs in the Control subcenters who did not use the KB system also reported that they would prefer to use the KB App moving forward. The average rating from ANMs in Treatment was 4/5 stars.

*Figure 21. App Ratings from ANMs using the KB App*

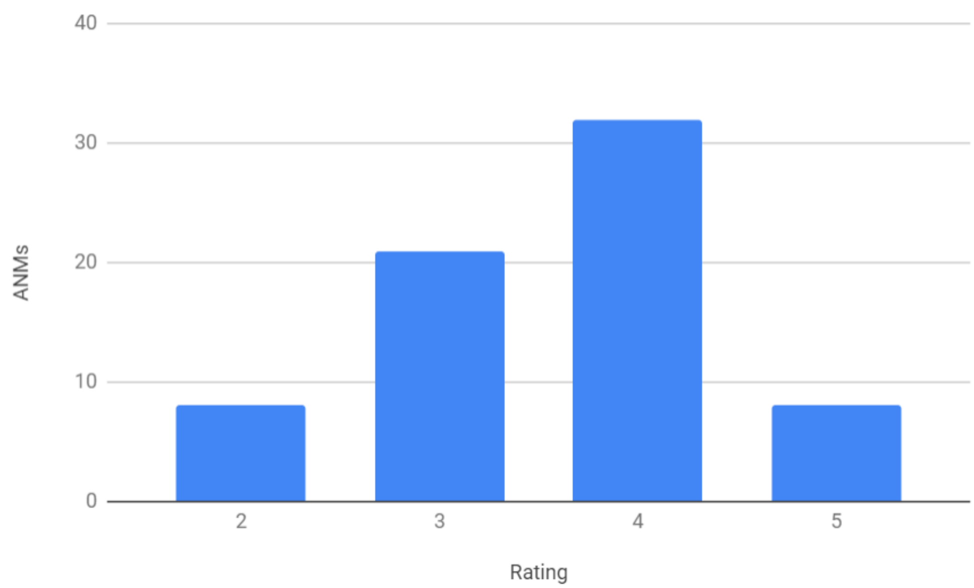
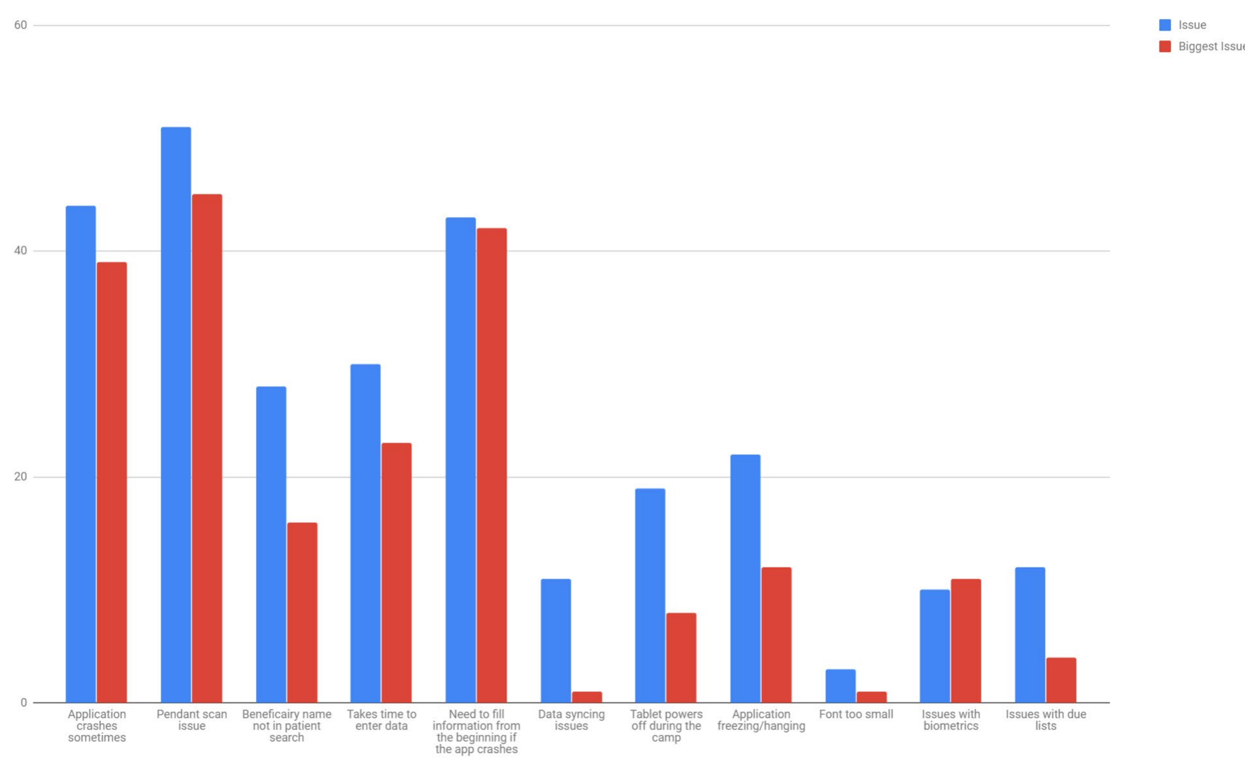


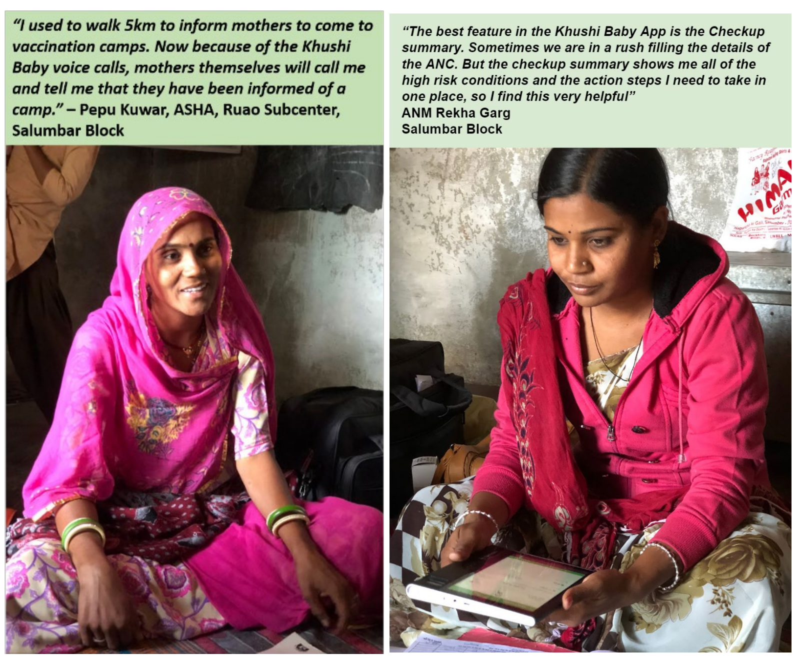

*Figure 22. Issues with the Khushi Baby App as reported by ANMs =*



The most prevalent issues reported by ANMs using the KB App included having to reenter data after an app crash (60.9%), having issues with scanning the pendant (65.2%), and issues with the application crashing (56.5%), and time required to enter data (33.3%). From our field observations common concerns included: charging of the tablet, app crashes, new registrations taking time, not having all peer ANMs use the app, differences in high risk classification, inability to edit the LMP date, and the fact that tablets had been changing hands frequently. Representative quotes from these ANMs can be found in the Appendix.

The majority of ASHAs reported that they found the system beneficial. In particular the voice call feature reduced their workload as they reported that mothers would come on their own, call them to check if the camp was occurring after receiving the Khushi Baby voice call, and even come before the camp started. They attributed the voice call to benefitting the antenatal care checkup coverage, the institutional delivery rate, and the full immunization rates in their communities. They reported that people liked the pendant and asked questions about where to get pendants that they as ASHAs would wear. They also appreciated how many of the due lists that they were responsible for tracking were autogenerated in the ANMs tablet. Representative quotes from these ASHAs can be found in the Appendix.

Feedback from health officials was largely positive regarding increased turn-up at camps and access to data, with specific references to WhatsApp and voice calls. Officials commented on how the system could be strengthened with direct integration into the PCTS backend, collection of additional data (e.g. vaccine logistics), uniform district-wide scale-up, additional content for WhatsApp group messaging and voice calls, and a focus on simplified, actionable views for the dashboard.



## Cost Analysis

We sought to assess the financial costs of the system (incurred and averted) to provide policy-makers with data to inform budgets for continued or scaled implementation. We used a back-of-the-envelope methodology developed by JPAL to list the ingredients-wise costs for the program over the entire study duration (July 2016 – July 2018). Incurred costs were determined from receipts of real expenditures as opposed to initially planned budgets and classified as upfront or recurring. Opportunity costs were included in the averted cost analysis. Where costs were left unmeasured, assumptions were made. We assumed no costs for targeting of beneficiaries to the intervention, given that implementation of such a system within established VHND camps would not trigger new targeting activities prior to implementation in a real-world sense, nor did we target beneficiaries in this study. We also assumed no additional costs to the participant (i.e. the mother/primary caregiver) outside what they would already occur for attending regularly scheduled VHND camps. This assumption is aligned with current government policies surrounding free provision of care to beneficiaries receiving services from VHND camps, irrespective of the data collection or community engagement methodologies employed. We assumed no additional cost to health workers for implementing the solution, given that the technology provision would not be borne by them, and that the use of the system (and trainings therein) would serve to complete their mandated scope of work currently being compensated by their government employer. Below Table 25 outlines the incurred costs (both upfront for project establishment and the recurring costs throughout the duration of the study), Tables 26-28 explain the assumed averted costs, and Table 29 summarizes the cost per beneficiary for the pilot and from the perspective of marginal costs moving forward, should the system be adopted by a government implementing body.

### *Table 20. Cost Ingredients for Cost Analysis*

| | **Base Year** | **Reporting Year** | **Notes / Assumptions** |
|---|---|---|---|
| | 2016 | 2018 | The costing period considered was from July 2016 to July 2018. Specific activities have time periods mentioned below. |
| **Program administration and staff costs** (July 2016 – July 2018) | ₹32,22,214 | $48,093 | Staff Costs: **$20651 (recurring)**<br>Includes: 25% time contribution for the COO, the Technical Program Manager, and the Field Data Manager; 100% time contribution from the Field Implementation Manager, Field Communication Lead, and Field Supervisor.<br><br>Program Administration Costs: **$27442**<br>Includes: Rent and utilities for headquarters **($11369, recurring)**, mobile phone talk-time, software subscriptions, and web server hosting subscriptions **($3379, recurring),** indirect costs and office expenses including laptops, chairs, desks, stationaries (**$12694, upfront)** |
| **Targeting costs** | | | All reproductive age women and mothers with children are already known to be eligible for the intervention which takes place at the VHND site. In a real implemention of this system, there would not be pre-intervention targeting of users. Just as was the case in this study, no additional costs were given to target beneficiaries towards VHND camps with the intervention |

| | Base Year | Reporting Year | Notes / Assumptions |
|---|---|---|---|
| **Staff Training** (February 2017 – July 2018) | ₹98,289 | $1,467 | This activityincludes costs for training the primary users of the intervention, primarily ANMs (87). No formal trainings were held for supporting staff (LHVs, DEOs, MOICs, BCMOs) Although supporting staff members were also visited and provided demos, these costs were included under the "monitoring costs" activity as they were performed by field monitors during their regular activities. |
| **Participant Training** | | | There were no associated costs with training the beneficiaries. |
| **Implementation and program material costs** (July 2016 – July 2018) | ₹1,05,55,448 (actual)<br><br>₹11,627,448 (with tablet) | $157,544 (actual)<br><br>$173,544 (with tablet cost) | *Technology costs: (funded by multiple agencies)*<br>- Tech development: **$63550** over the first 6 months **(upfront)**<br>- Tech maintenance: **$36000** over the next 18 months **(recurring)**<br><br>*Procurement costs:*<br>- NFC pendants: **$21792**; includes S&H and customs fees over the total study period **(upfront)**<br>- Tablets: $**0 (provided by IDEMIA's CSR);** In case of purchase, assume **$16000** for 80 total tablets at $200/unit based on recent negotiated rate with tablet provider Data Mini over the total study period **(upfront)**<br><br>*Implementation costs:* **$36202**<br>Includes: transportation and logistics costs related to initial deployment across 365 villages **($32789, upfront)** and data plans for ANMs to sync data from February 2017 to July 2018 (**$3233**, **(recurring)** |
| **User costs** | - | - | We assumed no cost to the user. There was no evidence that the user spent more time on the KB System at endline. |
| **Averted costs** (July 2016 – July 2018) | ₹3,56,335 (assumes no time savings)<br>₹72,34,773 (assumes time savings) | $5318 (assumes no time savings)<br><br>$107982 (assumes time savings) | See detailed description below in Tables 26-28 |
| **Monitoring costs** (July 2016 – July 2018) | ₹44,64,612 | $66,636 | This head includes includes salaries for field monitors and travel for field monitors over the study period. Field monitoring was a core component of the intervention. This head does not include costs related to baseline, midline, or endline surveys. |
| **Total Upfront Costs** | ₹98,37,275 | $146,825 | Includes office expenses for establishment, upfront costs for development, equipment and goods costs related to tablets and pendants, and initial costs of deployment in the field. |
| **Total Recurring Costs** | ₹47,73,951 | $71,253 | Includes program staff costs, rent and utilities costs, maintenance costs for the technology platform, costs of syncing, and monitoring costs. |
| **Total program cost** | ₹1,14,27,793 (assumes on time savings)<br><br>₹18,306,230 | **$273227** (assumes no time savings)<br><br>**$170564** (assumes time savings) | Total Program Costs = (Program Administration and staff costs + Targeting Costs + Staff Training Costs + Participant Training Costs + Implementation and Materials Costs + User Costs + Monitoring Costs) – Averted costs. In both cases, tablet costs are included. |

Averted costs are defined as costs that would have otherwise been borne by the payer if not for the intervention. The payer in this case is the district government, which receives funds from the state. We estimate the averted costs that come from three sources: increased immunizations rates (for pentavalent 1-3 and measles vaccines), decreased cases of moderate acute malnutrition, and time and supplies saved as seen in Tables 26-28 (See Appendix).

Averted costs from immunizations are estimated by multiplying the disease specific incidence in India by the vaccine attributable risk reduction by the increase in vaccination proportion in the treatment by the total costs of disease treatment and the process of seeking treatment. To estimate for longer term net benefits, the averted costs from disease treatment and the process of seeking treatment is multiplied by 1.67 in accordance with the results published by GAVI [9]. These estimates may be considered conservative given that the benefit from increased BCG and OPV was not accounted for due to lack of available costing data. Averted costs from the decreased rate of malnutrition were calculated by multiplying the difference in MAM incidence proportion by the number of children in the Treatment sample by aggregate cost of treating a child for MAM, based off data from a recent study in Mumbai slums [8]. This is likely a conservative estimate given that the broader economic benefits from a wellnourished child who is more likely to meet developmental and learning milestones were not accounted for in this estimate. Finally averted costs from time and supply savings were calculated using a series of assumptions. These time and supply savings were likely not realized during the study period due to lack of a state mandate to integrate Khushi Baby data directly from the treatment-randomized subcenters. However, we make estimates from our experience working with local stakeholders on the time saved in particular from monthly reporting tasks, currently on paper, which would be automated via the Khushi Baby application. Such averted costs can be relatively large in comparison to the health benefits, and therefore represent an optimistic or best-case scenario should the Khushi Baby platform be universally deployed

*Table 21. Aggregate Costs for Cost Analysis*

| | **Local (1 INR = $0.015)** | **USD (67 INR)** | **Notes** |
|---|---|---|---|
| **Total program cost** | ₹4,491,663 | **$273227** (assumes no time savings)<br><br>**$170564** (assumes time savings) | |
| **Number of total beneficiaries during the study period** | 19811 | Beneficiaries (mother and babies) | Total beneficiaries tracked for the Ante Natal Care, Post Natal Care, High Risk follow up and Immunization. Although the system was active between Feb 2017 and July 2018 (18 months) we conservatively assume that this total target achieved represents the actual target achieved over the 24-month total study duration, rather than extrapolating the beneficiary count to fit the total study timeline. |
| **Average cost per beneficiary over the full 24-month study period** | ₹924<br><br>₹577 | $13.79 (Assumes no averted costs from time savings or supplies)<br><br>$8.61 (assumes averted costs from time savings and supplies) | The average cost divides the total program cost as defined in Table 25 by the total number of beneficiaries served during the study duration. This includes both the upfront and recurrent costs during the study period. |
| **Marginal cost to add a beneficiary per year** | ₹376.54 | $4.47 | NFC Pendant, customs, transportation = $0.80 per year assuming 1 pendant for the mother during pregnancy and 1 pendant for the child during infancy<br><br>Voice calls, 7 reminder voice calls per year at $0.03 per voice call = $0.21<br><br>Tablet ammortized over 1000 patients per ANM in a 2-year span = $0.04 in total or $0.02 per year<br><br>Tech maintenance = $36,000/18 monthss * 12 months = $1.21 per year<br><br>Monitoring: $66,366 over 18 months for 19,811 beneficiaries = $2.23 per beneficiary per year |

| | | | |
|---|---|---|---|
| | | | *Cost of data plans for ANMs and data hosting was not included given the availability of existing government cellular networks and servers which can be leveraged for the purpose if/when the system is scaled. |

Using the JPAL's Basic Cost Collection Template, it was determined that the average cost per beneficiary during the study period (24 months) was $13.79 (or $6.90 per year) in the conservative scenario with no assumed averted costs from time savings and $8.61 over the two-year period ($4.30 per year) in the scenario with included averted costs from projected time savings. The marginal cost to add a beneficiary per year to the system was $4.47 assuming the conservative case with no averted costs, which is our preferred estimate moving forward as it relies on the fewest assumptions. If averted costs are included without projected time savings, then the marginal cost to add a beneficiary per year is estimated to be $3.73 less ($0.74). If averted costs are included with projected time savings, then the marginal cost to add a benefeficiary per year is estimated to be $5.45 less ($0.98 net profit).Notably tablet costs were included in all estimates, even though not part of actual costs borne by the team. At scale, existing tablets for which an investment has already been made could be repurposed for this use case. The yearly averted cost calculation also makes assumptions, which were not directly proven in this study.

### *Cost Effectiveness Analysis*

We used a cost effectiveness analysis method described by Borkum et al [10] in their similar, but smaller study of an mHealth application for community health workers in Uttar Pradesh Bihar under a similar randomized controlled design. They defined cost effectiveness as the cost per beneficiary per year for a 1 absolute percentage point increase in the outcome compared to the baseline. And due to multiple outcomes being evaluated they suggested their estimate was a high-end cost per utility estimate compared to a scenario in which all resources of the intervention were vertically focused on a single outcome as opposed to a set of multiple outcomes. We took forward the same assumption given a) that our intervention's combined focus on MCH made it arbitrary to separate costs for any particular outcome (immunization vs. nutrition for example) and b) that it was the most conservative to assume that all the resources (and therefore costs) were focused on the primary outcome of interest – immunization.

Because at baseline recall data were collected on immunization status, a difference-in-difference was calculated for an unadjusted mean average difference in absolute percentage points as a result from the baseline of 12.2 percentage points (95%CI 3.5-20.9%). Using this metric, the cost effectiveness analysis for this unadjusted estimate could be interpreted as $4.47 per additional beneficiary divided by 12.2 percentage point absolute increase in full immunization rate or $0.37 (95% CI $0.21 - $1.28) per beneficiary per year per 1% absolute increase in full immunization rate.

Bokrum et al used adjusted differences so a direct comparison is not appropriate. But for sake of reference, they found a $1.02 cost per beneficiary per year per 1 absolute percentage point increase in mothers receiving full tetatanus vaccine coverage, which serves as the closest proxy indicator. Their study did not show a significant yield for improvement of full infant immunization.

Using this study's adjusted effect estimates, the cost-effectiveness can alternatively be calculated by looking at cost per percent change in relative risk. In this case, the cost effectiveness would be $4.47 per additional beneficiary divided by 66% (95% CI 23%-124%) relative increase in full infant immunization or $0.68 (95%CI $0.36-$1.94) per 10% relative increase in the likelihood of full infant immunization

To determine the cost per additional child fully immunized, we divided the total program cost ($273227 in the case without time savings and $170564 in the case with time savings) by the additional number of infants in the treatment arm fully immunized compared to the control arm (12.2 percentage points or 140 infants). This results in a cost per additional child fully immunized in the intial roll-out phase of $1950(95% CI $1150-$6860) in the case where time and supply savings are ignored and $1220 (95% CI $710-$4280) in the case that assumes time and supply savings. These are conservative estimates which assume that the entire benefit of the intervention was derived from the increase in immunization. This cost effectiveness estimate does not include the overlapping cost for program activities that lead to improvements in infant malnutrition or for the services received by out-of-sample mothers who were impacted by the intervention during the study period but not considered for the final outcome.

To calculate the average cost per additional vaccine we first found the difference in average vaccinations between the treatment and control group using recall data as 0.52 (95% CI 0.34–0.70). This difference was multiplied by the total Endline Treatment sample size to determine the number of total additional vaccines received: 600 (95% CI 390 – 810). The total program cost was divided by the total additional vaccines received to obtain the cost per additional vaccination: $460 (95% CI $340-$700) in the case with no time savings included and $280 (95% CI $210-$440) with time savings included.

The estimates above serve as a rough calculation of the cost effectiveness for the pilot stage roll-out. After initial deployment, the cost of subsequent beneficiary enrollment drops from $6.90 to $4.47 due to gains from efficiencies of scale in technical management and operations balanced by increased admin costs.

# Discussion

## Challenges and Lessons Learned

There were significant challenges faced in the implementation of this project from financial, political, technical, operational, and intra-team domains.

This project was initially funded by a grant from 3ie, which was circumscribed to impact evaluation. At the time the grant was received however, the intervention-to-be-tested had yet to be developed. We severely underestimated the costs of technical development of our 2.0 system (KB App and Dashboard) which would cover the MCH tracking requirements laid forth by the government heatlh officials. The significant costs related to technical development ended up being over $100K over what we had projected. These costs were coupled with delays in receiving anticipated funding from two granting agencies. For three months key field staff and core team members went without a salary in the summer of 2017 while we had to seek bridge loans until our next funding tranche was resolved. Access to capital was also limited by the fact that as an Indian non-profit organization, there have been barriers to receiving funds from international agencies abroad. As an independent entity and early stage start-up, we were not successful in convincing the local district to put forth a financial stake. We should have either requested more money or planned for a smaller trial. But a smaller trial would also risk not having adequate analytical power due to the necessary clustering to avoid the health worker from using both the KB system and the paper intervention in different camps. We found a way forward with timely receipt of other grants to fill the deficits from the tranche delays and by having the core team defer salary towards technical development. Beyond technical development, this grant also went towards establishing an HQ and hiring our first full-time staff. Due to budget limitations, the staff initially comprised of four members who were managing a team of over 40 surveyors during the baseline survey exercise. Lack of financial stability early on certainly affected the execution of the project.

Multiple political hurdles were overcome in this impact evaluation. It took over a year of engagement with the CMHO to convince him that the organization was serious and here to stay as opposed to other pilots which had worked with the local government in the past. In one case, our COO had to track down the CMHO on the highway after almost daily visits to the chamber seeking approval during which the CMHO did not have time to fully see the system. The CMHO initially was insistent that the evaluation be conducted using a quasi-experimental design with all ANMs of one block receiving the intervention to avoid administrative burdens from the randomization. After much persuasion, the CMHO acquiesced to the current study design. Even with the CMHO's support, there was no official written mandate from the District or MOU that established that ANMs had to adhere to the Khushi Baby study or intervention protocols. This proved to be especially challenging with respect to the ANMs in Lasadiya who were increasingly resistant to adopting the new system. The local health officer there also had political clout and turned a blind eye despite multiple visits in which the team informed him of poor acceptance by ANMs in his area. We were ultimately unable to change the minds of some 5-10 ANMs. We also failed to regularly engage the health officials and supervisors at various levels. And with officials changing every few months, we did not have the policy manager in place to maintain relations with the new apointees. We have now hired someone to fill the position and have activated our Field Monitors and Core Team to regularly track and participate in block and district level meetings with local officials.

One way our lack of engagement translated into the study was by means of the health officials not using the dashboard. This lack of use may be most attributed to the fact that our system was not rolled out through an entire catchment area, thereby only providing partial data for action. Moreover, the health officials and ANMs had to submit data according to the state mandated PCTS system. And because of this, ANMs were still bound to the double-work of filling paper forms both at the camp and on a monthly basis. The way we handled these challenges was by adapting the reports of the dashboard into a format accessible on WhatsApp. This proved to get various health workers engaged, although the quality of responses towards high risk patients still left room for improvement among ANMs and supervisors alike. For double-work we introduced features like the patient search and summary which allows the ANM to have quick views of all the beneficiaries they had seen in a month.

There were many substantial technical hurdles at play, both in the development and maintenance phases of the project. The system was developed without the guidance of a professional designer by the core team and as a result many UI/UX elements were not optimized. Certain features of the technical architecture were also particularly weak including handling of backend database joining, optimization of backend memory usage, and handling of NFC reading and writing. We were unfortunately provided with over 80 tablets that had unreliable hardware components which necessitated replacement. Four months after starting we discovered data corruption issues due to the faulty hardware that we had replaced, which had us backtracking to solve the issue for over 200 beneficiaries with corrupted data. Some of the technical errors may have been averted had we been able to earlier hire in-house developers who were closer to the field issues and who could more readily communicate with our testers. We later developed systematic processes for documenting our testing of each build and reached stability after a year of iteration in the field. Still we have persistent issues with some of the hardware-specific software for NFC handling. We look to resolve these issues in our next 3.0 system with new hardware and a newly built platform, which will also be one carefully co-designed with the user, led by our newly hired Lead Designer.

Operationally this was a very challenging project for several reasons. With our HQ centralized in Udaipur we were still having to oversee a radius of up to 150 km, which included the hilly terrain of the Aravali Range. Maintaining regular communication with our own field monitor team took time as we first had to identify appropriate network points near each of the nearly 600 villages in the study. Field monitors and the core team were heavily involved in coordinating syncing of data from the field at these network points as well as coordinating replacement of faulty tablets. If a tablet was causing issues, getting data to understand the root cause required going to great lengths to retrieve that tablet and bring it to the HQ. The team also had to battle through the elements in the monsoon season when bridges were broken and in the hot dry season when temperatures would climb over 45 C degrees for most of the day.

Our team was responsible for designing, working closely with developers, testing, training, deploying, providing in-field support, monitoring and evaluating the project, and interacting with stakeholders. With limited funds, we did not always have the full capacity to address these technical issues in-house. With limited staff in the early stages, there were periods when staff were severely overworked and there were concerns that several members may leave the organization for new jobs. Many of the team members volunteering globally also became occupied with various other obligations and could no longer be held to the same level of accountability as when the project had started. Despite these challenges our organization grew from four full-time members at the study start to over 20 full-time staff today, including new technical talent in data analysis, design, and development.

Future challenges can be mitigated with proper planning. Financially, we would like to ensure an 18-month runway with out funders and to schedule large field deployments or procurements after receiving a new tranche of funding for the allocated activity. Conservative impact estimates and timelines will be made to donors to avoid no-cost grant extensions. Technically, we will recruit in-house talent. A technical program manager will use project management tools for agile application design. Weekly code and progress reviews will take place by a senior technical lead. A hardware senior person of contact will be appointed by the vendor and deployed to the field site during the launch phase.. Prior to development, a rigorous design process will take place – defining the information architecture for the overall application, preparing high-fidelity visual mockups of each screen, and gaining feedback from ANMs. Politically, we will increase our engagement with the District and State level by having biweekly meetings with concerned decision-makers and policy influencers and by participating in working groups for health information system integration

## Substantive Findings from Impact Evaluation

With respect to Khushi Baby system functionality, there were initial hardware and software challenges faced, during which all tablets had to be replaced. By endline, among the 87 ANMs, the last 90 days shows 93.7% crash-free sessions and 75.4% crash free users, showing reasonable stability with room for improvement, and the median time to receive health updates was 3.45 hours (compared to a maximum time of 30 days as mandated by the existing paper-based process). Retention of the health record was greater for the pendant (82.4%) when compared to the MAMTA card (74.8%) at Midline ($p = 0.0031$) but was not significantly different at Endline (78.3% vs. 77.9%). Higher retenton of the heatlh record at any time point is meaningful for ensuring proper receipt of services at the camp site.

For those children who did retain their health records, data completeness (98.70% vs. 77.11%, $p < 0.00001$) and consistency (98.1 vs. 81.0, $p<0.00001$) were increased in among those who had the Khushi Baby pendant. By design of the KB system, records cannot be created without passing certain validations, so the high rate of data completeness was expected for the KB system, provided it functioned according to plan. With respect to the validity of the data entered on the MAMTA card, 27% of the fields were correctly recorded (as opposed to skipped or falsely entered) for children, and 49.8% for mothers. There were no significant differences in the child or antenatal care checkup processes conducted or the rates of correct data entry on the MAMTA card when comparing Treatment and Control camps. Altogether, these findings suggest the Khushi Baby system's data had higher overall quality, but there may be room for improvement with respect to checkup and data-filling processes.

The data collected from the Khushi Baby system was used for new forms of substantive action in the field, particularly through WhatsApp groups, automated voice call reminders, in-person voice call outreach, and high-risk child follow-up house visits. On a weekly basis over 20 WhatsApp messages are exchanged and over 1100 voice call reminders are automatically sent to mothers on average. While health officials did not use the dashboard as intended, reports from the dashboard were sent by the Khushi Baby team members on the WhatsApp groups where health officials and health workers responded to specific high-risk patient follow-ups.

With evidence for new data-driven actions from the Khushi Baby system, we looked to investigate the impact on health behaviour outcomes, principally towards immunizations. Knowing that the data quality of the common measurement tools for full immunization were subject to incomplete and false entry we sought multiple methods of corroborating the vaccination status: both from recall and by examining the MAMTA card. We found that mothers randomized to the Treatment subcenters were 1.66 times (95% CI 1.23-2.24) more likely to have their children fully vaccinated when compared to mothers from Control subcenters, irrespective of receiving the intervention by way of recall and after adjusting for potential confounders under a conservative Intent-to-Treat framework. This result remained consistent when looking at MAMTA cards in Treatment and Control groups (OR 1.35, 95% CI 1.10-1.67), and when considering full immunization by 9 months for both methods and improvmeents in pentavalent 1-3 coverage. There was a 12.2 percentage point difference-in-difference in absolute full infant immunization rate between Treatment and Control from baseline to Endline. The benefit of the intervention was not partial to any subgroup on the basis of SES quartile, distance to camp, or baseline maternal child health awareness. Factors independently associated with a higher full immunization rate included number of reported KB calls received, SES score, ASHA visit frequency, ANM proximity to the PHC, whether a family member prevented vaccination, and geographic block, with Lasadiya performing worst among its peers, among covariates. These results are particularly noteworthy even though the research cohort only began to receive voice calls during the infancy period. The full potential of receiving the intervention during antenatal care, which is the standard for non-sample mothers, was not captured in this study.

With improvements in full immunization outcomes, we looked to investigate how health behaviors may have contributed to improved health outcomes. We found that infants of mothers randomized to Treatment subcenters were 26.4% (95% CI 6.3-42.3) less lilely to have moderate acute malnutrition (MUAC between 11.5-12.5 cm) compared to children of mothers randomized to Control subcenters, irrespective of receiving the intervention, and after adjusting for confounders. This result was particularly significant because unlike data reported on the MAMTA card or by recall from the mother, the data for the outcome were measured directly by using a mid-upper-arm circumference tape. There were no significant differences found for rates of severe acute malnutrition, 24-hour hospitalizations of the infant, hospitalizations due to diarrhea, hospitalizations due to suspected pneumonia, or overall infant mortality, although the study was not powered for these outcomes.

To understand the underlying drivers for the observed health behaviour outcomes, we asked mothers at endline about their attitudes and perceptions towards their recent health journey. Notably, more mothers in Treatment felt empowered to make decisions on their health ($p = 0.00052$), whereas proportional more mothers in Control reported feeling more comfort with receiving health care ($p < 0.00001$). However, more mothers in Treatment felt comfort towards exclusive breastfeeding practices ($p<0.00001$). This may have consequence on the lower rates of moderate acute malnutrition found above. As expected, a higher proportion of mothers attributed their increase in awareness and their husband's increase in awareness to mobile messages they received during the intervention period ($p < 0.00001$). Mothers report that they receive the messages, they listen to the messages, and they understand the messages. They also have been asked by others in the village about where they received their pendant. They feel empowered when the ANM takes their biometric to access their health history.

We were also concerned with how this system would affect our users. Feedback from our ANMs, the key users of the KB App, has been positive. 62/69 (90%) of ANMs who used the app reported that they would prefer to continue to use the system moving forward. 68/77 (88%) of ANMs in the Control subcenters who did not use the KB system also reported that they would prefer to use the KB App moving forward. The average rating from ANMs in Treatment was 4/5 stars. ANMs cited benefits including ease of identifying high risk patients, automatic generation of due lists, time saved at the camp, and noticeable improvement in camp turn-up from voice calls. ASHAs were indirect beneficiaries of the system. Although they did not use the app, they strongly reported that due to Khushi Baby voice calls that they had seen improvement in uptake of antenatal care, hospital delivery, and uptake of child immunization. Before Khushi Baby they would have to go repeatedly to certain households to remind them for upcoming camps. Now instead mothers and pregnant women are receiving the Khushi Baby call, calling the ASHA to confirm the camp the next day, and coming themselves, in some cases, before the camp even starts.

Health officails and supervisors strongly appreciated the system, particulary the WhatsApp groups through which they reported that they saw an improvement in how the ANM was addressing high risk and drop out beneficiaries. There were calls to have the system expanded and integrated into the Rajasthan State Ministry of Health's PCTS database, by the Chief Medical Health Officer of Udaipur, Dr. Sanjeev Tak, who appreciated how individual elements of the Khushi Baby system from the pendant to the application to the high-risk reports, voice calls and WhatsApp group fit into a theory of change for health outcomes. Altogether, the findings above represent strong evidence from a large, randomized, prospective trial, that even without the full scope of the Khushi Baby application in play and despite challenges, there were significant improvements in data quality, health behaviour outcomes, health outcomes supported by health worker and health official satisfaction.

### Comparisons to Existing Literature.

The study was conceived as an extension of a prior randomized controlled trial by the Khushi Baby team in 96 villages of Udaipur with a partner NGO, Seva Mandir, for which an under-powered sample failed to show significant differences in timely immunization coverage through DTP3, but which did show significant increases in the levels of discussion concerning the pendant vs. the MAMTA card [3]. This impact evaluation here is unique in that it is the first of its kind to test a combination several novel components for community engagement including but not limited to the wearable[15], culturally-symbolic, NFC-based pendant[16] and dialect specific voice calls[17] for MCH education against the status quo control as the gold-standard. It is the first of its kind large scale randomized trial for a mHealth intervention that considers the ANM as its key user, and the largest of its kind ever in the Udaipur region to our knowledge. Other related studies have looked at providing ASHAs with mobile applications to improve uptake of essential MCH services in India [10, 11] and in redesigning the health card [12].The findings here support the literature from other mHealth interventions in the LMICs, particularly around the use of mobile reminders (many of which have been SMS based to date) to improve immunization timeliness and coverage and other health behaviour outcomes [13-19]. This impact evaluation corroborates findings from many other studies which outline several key covariates associated with uptake of essential MCH services including socioeconomic status, distance from health camp, opinions of members of the

---

[15] Alma Sana is a wearable silicon bracelet with punch-out shapes that represent vaccines completed
[16] Marcus et al used an RFID based ankle bracelet for pneumonia tracking in Pakistan for infants coming into a health care facility
[17] Mobile Alliance for Maternal Action ("MAMA"), and mMitra – ARMANN are two voice based reminder platforms for maternal child health for LMICs

household, along with qualities of the health worker and local health volunteer which comprise both demand-side and supply-side factors [20-27].

## Strengths and Limitations of the Study

There were several limitations. First the study was un-blinded, which makes it difficult to distinguish the success from the real intervention against any placebo intervention. That being said, the underlying mechanisms in the theory of change were investigated to build evidence for or against a causal link between intervention and outcomes. Another limitation of this study is that the intervention was not rolled out at once leading to a possibility of time-based confounding, and that the intervention evolved with time. Also, the delay between baseline enrollment and intervention roll-out may have led to selection bias as the eligible infants were only considered from February 2017 onwards in Treatment whereas they were eligible in Control from as early as October 2016. We ultimately did not find evidence to support temporal confounding affecting the main results of the trial. While some may consider the degree of monitoring as playing a confounding role on the impact of the platform, we would instead argue for monitoring as a central component to the complete system. The study strengths include the prioritization of data quality as an outcome, the nearly 10,000 interviews conducted with mothers, ANMs, health supervisors, and health officials at every level to include quantitative and qualitative feedback, and the randomized prospective design to account for measured and unmeasured sources of confounding and to establish causal chronicity of events, the high level of follow-up at Endline (88%), the adjustment for potential contimanating interventions such as the Mission Indradhanush door-to-door immunization scheme, and the corroborations of multiple sources of data for outcomes with highly granular information about intervention uptake such as duration of voice call listened to by the beneficiary.

## Implications for Policy Makers

The findings of this impact evaluation demonstrate the highest grade of evidence for the impact of the Khushi Baby system for improving: data retention, data timeliness, data quality, data for action, health behaviour outcomes, and health outcomes for infants from poor households in rural settings. Specifically, the data showed Khushi Baby increased immunization rates by 12 percentage points decreased acute malnutrition rates by 4 percentage points, improved data completeness and data consistency by nearly 20 percentage points, and reduced the time to acquire data to a median of just under 4 hours with an average cost of $4.47 or just about 400 INR per beneficiary per year, with a 4/5 star approval from ANMs.

Khushi Baby is one of several mHealth applications for MCH in India. Antara's AAA Platform for ANMs, ASHAs, and Anganwaadi workers is another Android-based tablet app. The Central Ministry of Health and Family Welfare's ANMOL is another which was tested with over 11,000 ANMs in Andra Pradesh. Recently, Dimagi announced a partnership with the Ministry of Woman and Child Development to scale up ICDS-CAS, mobile based data collection app for nearly 100,000 anganwaadi workers. The government of Rajasthan along with other states across the country are seeking bulk orders of tablets to empower each frontline health worker. Recently, implementation of a PCTS mobile app by the Rajasthan MOHFW has come across resistance from ANMs who perceive the task as a form of extra data entry. Rajasthan also faced resistance with an earlier attempt to procure and distribute Micromax tablets to bring the mobile app E-janswasthya to scale, but had recent success in 2019 in using a version of the system when ASHAs were able to self-procure their own mobile phones without incentives to complete a digital family health survey in two blocks.

In many cases, these applications have been considered for scale-up without a base of evidence to prove stability, acceptability, or effectiveness at the local level. In other cases, pilot projects have been abandoned after large scale experimental trials, despite promising results, as was the case with ICT-CCS in Bihar in 2015. Few examples have successfully scaled such as Gujarat's TECHO+ app, which combined strong randomized results with state and central-level political will for scale-up across multiple health worker cadres in the state. With Digital India, Smart City Initiatives, AADHAAR, and the recent Univeral Payment Interface, digitization will only continue to expand throughout India, even towards the last mile. The question then becomes not if a platform is worth replacing existing paper-based systems, but rather which platform should be considered for scale-up. Given the significant costs associated with training and replacing an existing system that has for so long worked on paper, assumptions on any digital platforms effectiveness ought to be evaluated thorouighly before crores of investment are spent on tablets, building software, and training of health workers.

Khushi Baby's unique value proposition begins with rigorous evidence and over four years of experience working and overcoming significant barriers at the last mile. The value centers around the accountability of a decentralized digital health record and the automated dialect-specific voice call for engagement. Most importantly, Khushi Baby's solution is offered as an end-to-end solution including design, development, deployment, mobilization, and monitoring and evaluation in an iterative loop.

## Generalisability of Findings

The findings of this impact evaluation may be generalized to the full five administrative blocks of Gogunda, Jhadol, Lasadiya, Salumbar, and Sarada. The subgroup analysis suggests that the intervention is not partial to groups based on socioeconomic status, health awareness, or distance to health camp. Notably, this evidence has been used to drive policy of the Udaipur District government, and the Khushi Baby system will be scaled to cover these blocks universally by early 2019. These findings may also be relevant for other areas across India, which share cultural beliefs in the black thread, and where populations of a similar demographics (living on less than 1000 INR monthly, with high access to mobile phones, low literacy, low baseline MCH awareness, with predominantly agricultural labor as the primary source of income, within 5km of a health sessions camp) having low baseline antenatal care and full immunization rates (less than 40%) also reside. The findings however should not be generalized to other interventions that only capture certain elements of the intervention or which do not cater to the continuous improvement of the solution life-cycle.

## Online Appendix

https://www.3ieimpact.org/sites/default/files/2020-06/TW10.1078-Online-appendix.pdf

## References


1. *Pregnancy Child Tracking and Health Services Management System (PCTS) Gap Analysis Rajasthan*. 2015, Columbia Global Centers: Mumbai.
2. Sherpa, D., *A Study on Developing an mHealth Solution as an Alternate Recording and Reporting System for Auxiliary Nurse MIdwives in Udaipur District of Rajasthan*. 2016, Indian Institute of Health Management Research: Jaipur.
3. Nagar, R., et al., *A cluster randomized trial to determine the effectiveness of a novel, digital pendant and voice reminder platform on increasing infant immunization adherence in rural Udaipur, India.* Vaccine, 2018. **36**(44): p. 6567-6577.
4. Zhang, Z. and A. Castelló, *Principal components analysis in clinical studies.* Annals of translational medicine, 2017. **5**(17): p. 351-351.
5. Wahi, M., *Healthcare Analytics: Regression in R*. 2017, Lynda.
6. Education, U.I.f.D.R.a. *MIXED EFFECTS LOGISTIC REGRESSION | R DATA ANALYSIS EXAMPLES*. [cited 2018; Available from: https://stats.idre.ucla.edu/r/dae/mixed-effects-logistic-regression/.
7. Megiddo, I., et al., *Analysis of the Universal Immunization Programme and introduction of a rotavirus vaccine in India with IndiaSim.* Vaccine, 2014. **32 Suppl 1**: p. A151-61.
8. Goudet, S., et al., *Cost effectiveness of a community based prevention and treatment of acute malnutrition programme in Mumbai slums, India.* PloS one, 2018. **13**(11): p. e0205688-e0205688.
9. Ozawa, S., et al., *Return On Investment From Childhood Immunization In Low- And Middle-Income Countries, 2011–20.* Health Affairs, 2016. **35**(2): p. 199-207.
10. Evan Borkum, A.S., Swetha Sridharan, et al., *Evaluation of the Information and Communication Technology (ICT) Continuum of Care Services (CCS) Intervention in Bihar*. 2015, Mathematica Policy Research.
11. Prinja, S., et al., *Impact of m-health application used by community health volunteers on improving utilisation of maternal, new-born and child health care services in a rural area of Uttar Pradesh, India.* Trop Med Int Health, 2017. **22**(7): p. 895-907.
12. Usman, H.R., et al., *Randomized controlled trial to improve childhood immunization adherence in rural Pakistan: redesigned immunization card and maternal education.* Trop Med Int Health, 2011. **16**(3): p. 334-42.
13. Gibson, D., *The readiness, need for, and effect of mhealth interventions to improve immunization timeliness and coverage in rural western Kenya*, in *International Health*. 2014, Johns Hopkins University
14. Bangure, D., et al., *Effectiveness of short message services reminder on childhood immunization programme in Kadoma, Zimbabwe - a randomized controlled trial, 2013.* BMC Public Health, 2015. **15**(1): p. 137.
15. Chen, L., et al., *Effectiveness of a smart phone app on improving immunization of children in rural Sichuan Province, China: study protocol for a paired cluster randomized controlled trial.* BMC public health, 2014. **14**: p. 262-262.
16. LeFevre, A., et al., *Forecasting the Value for Money of Mobile Maternal Health Information Messages on Improving Utilization of Maternal and Child Health Services in Gauteng, South Africa: Cost-Effectiveness Analysis.* JMIR Mhealth Uhealth, 2018. **6**(7): p. e153.
17. Lester, R.T., et al., *Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial.* Lancet, 2010. **376**(9755): p. 1838-45.

18. Uddin, M.J., et al., *Use of mobile phones for improving vaccination coverage among children living in rural hard-to-reach areas and urban streets of Bangladesh.* Vaccine, 2016. **34**(2): p. 276-283.
19. Oyo-Ita, A., et al., *Interventions for improving coverage of childhood immunisation in low- and middle-income countries.* Cochrane Database Syst Rev, 2016. **7**: p. Cd008145.
20. Vora, K.S., S.A. Koblinsky, and M.A. Koblinsky, *Predictors of maternal health services utilization by poor, rural women: a comparative study in Indian States of Gujarat and Tamil Nadu.* Journal of health, population, and nutrition, 2015. **33**: p. 9-9.
21. Lakew, Y., A. Bekele, and S. Biadgilign, *Factors influencing full immunization coverage among 12–23 months of age children in Ethiopia: evidence from the national demographic and health survey in 2011.* BMC Public Health, 2015. **15**(1): p. 728.
22. Banerjee, A.V., et al., *Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives.* BMJ, 2010. **340**.
23. De, P. and B.N. Bhattacharya, *Determinants of Child Immunization in Fourless-Developed States of North India.* Journal of Child Health Care, 2002. **6**(1): p. 34-50.
24. Fineberg, H.V., *The paradox of disease prevention: Celebrated in principle, resisted in practice.* JAMA, 2013. **310**(1): p. 85-90.
25. Group, S.W., *Report of the Sage Working Group on Vaccine Hesitancy*. 2014, WHO.
26. UNICEF, *National Factsheet: Coverage Evaluation Survey*. 2009.
27. Usman, H.R., et al., *Determinants of third dose of diphtheria-tetanus-pertussis (DTP) completion among children who received DTP1 at rural immunization centres in Pakistan: a cohort study.* Trop Med Int Health, 2010. **15**(1): p. 140-7.