ePact
Strengthening evaluation effectiveness and impact

UKaid
from the British people

External evaluation of mobile phone technology based nutrition and agriculture advisory services in Africa and South Asia

Mobile phones, nutrition, and health in Tanzania: Quantitative Baseline Report

Daniel Gilligan, Melissa Hidrobo, Natasha Ledlie, and Giordano Palloni
International Food Policy Research Institute

Final version –12 July 2018

Oxford Policy Management | itad | IDS Institute of Development Studies | Gamos | IFPRI International Food Policy Research Institute

e-Pact is a consortium led by Oxford Policy Management and co-managed with Itad

Acknowledgements

The International Food Policy Research Institute (IFPRI) thanks all of the households in Iringa, Tanzania that agreed to participate in this research project. We are also grateful to the Oxford Policy Management Tanzania team led by Andreas Kutka, Katie Mcintosh, and Lindsey Roots, for leading the data collection. We would additionally like to thank the Groupe Spéciale Mobile Association (GSMA), Cardno, the mHealth Tanzania Public-Private Partnership, the Tanzania Food and Nutrition Centre, the Tanzanian Ministry of Health, Community Development, Gender, Elderly and Children, Community Development, Gender, Elderly and Children, and the Global Alliance for Improved Nutrition (GAIN) teams for their ongoing cooperation and support for the quantitative component of the overall impact evaluation.

The IFPRI research team benefitted enormously from our ongoing partnership with the Institute of Development Studies (IDS) which oversees the qualitative component of the evaluation and the GAMOS team conducting the business model portion of the evaluation. Both teams have provided invaluable input and feedback as we designed the evaluation, developed tools, and composed this report. Finally, we would also like to thank the Department for International Development (DFID) and Oxford Policy Management (OPM) teams for their continued support and contributions to the project.

Disclaimer

This material has been funded by UK aid from the UK government; however, the views expressed do not necessarily reflect the UK government's official policies.

This project is being led by the Institute of Development Studies (IDS) together with GAMOS and the International Food Policy Research Institute (IFPRI) as part of the e-Pact consortium led by Oxford Policy Management (OPM) co-managed with Itad. The IDS project manager is Jessica Gordon [j.gordon@ids.ac.uk]. The report authors are Daniel Gilligan, Melissa Hidrobo, Natasha Ledlie, and Giordano Palloni (IFPRI). For further information please contact C.Gee@cgiar.org.

The contact point for the client is Louise Horner [l-horner@dfid.gov.uk]. The client reference number for the project is PO6420.

Data Management

All intellectual property rights in any materials produced from the evaluation (including publication of research findings and any other reports and data) remain the property of IDS and associated sub-contracted collaborators. IDS and all sub-contracted partners undertaking data collection have specific arrangements in place for handling data generated from the project in accordance with the Data Protection Act (1998) which includes the processing and storage of any sensitive personal data and maintenance of privacy. DFID has unlimited access to any material produced from the evaluation. In order to promote use and uptake of the evaluation findings and in line with DFID's Enhance and Open Access Policy, the evaluation team is committed to ensuring all major report outputs and associated data generated from this project are made publicly available in an accessible format. Following approval of the report from DFID, the baseline report will be made available on IFPRI's and IDS's websites. All datasets will also be made available within 12 months of final data collection on IFPRI's page at the Harvard Dataverse web site: https://dataverse.harvard.edu/dataverse/IFPRI.

This project is being led by the Institute of Development Studies (IDS) together with GAMOS and the International Food Policy Research Institute (IFPRI) as part of the e-Pact consortium led by Oxford Policy Management (OPM) co-managed with Itad. The IDS project manager is Jessica Gordon [j.gordon@ids.ac.uk]. The report authors are Daniel Gilligan, Melissa Hidrobo, Natasha Ledlie, and Giordano Palloni (IFPRI). For further information please contact C.Gee@cgiar.org.

The contact point for the client is Louise Horner [l-horner@dfid.gov.uk]. The client reference number for the project is PO6420.

e-Pact
Level 3, Clarendon House
52 Cornmarket Street
Oxford OX1 3HJ
United Kingdom
Tel +44 (0) 1865 207300
Fax +44 (0) 1865 207301
Email admin@opml.co.uk
Website www.opml.co.uk

Executive summary

The mNutrition Intervention in Tanzania

mNutrition is a five-year global initiative supported by the Department for International Development (DFID), organised by Groupe Spéciale Mobile Association (GSMA), and implemented by in-country service management organisations in cooperation with mobile network operators (MNOs) to use mobile technology to improve the health and nutritional status of children and adults in low-income countries around the world. mNutrition is implemented through existing mAgri and mHealth programmes in 12 countries throughout sub-Saharan Africa and South Asia. The nutrition content of the programmes aims to promote behaviour change around key farming practices and dietary and child feeding practices that are likely to result in improved nutritional health within a household.

DFID has committed to conducting an independent evaluation of mNutrition. Given the scale of the mNutrition programme the decision was made to select two countries for inclusion in the evaluation: the mHealth programme in Tanzania and mAgri programme in Ghana. The mNutrition intervention that is the focus of the evaluation in Tanzania and this report is an integrated programme that combines an existing SMS-based health communication campaign that targets pregnant women and mothers of young children known as “Healthy Pregnancy, Healthy Baby” with approximately 120 nutrition focused SMS messages in Swahili.

The combined programme will send out SMS messages timed to the life cycle of mothers and children. It is expected that the programme will provide relevant information for pregnant women and children up to the age of five and that the information will affect beliefs and behaviours in key nutrition related areas—including Infant and Young Child Feeding (IYCF) and women’s dietary diversity—for recipient households.

Baseline data suggest there is scope for the sample to benefit from the mNutrition programme. Access to and use of mobile phones is high—90.7% of interviewed men and 58.9% of interviewed women own a mobile phone, while 83.5% of men and 62.4% of women used a mobile phone to receive an SMS message during the previous fourteen days. Further, 82.0% of interviewed women indicated that they would trust health or nutrition information received through automated text messages from an NGO, private organization, religious or voluntary organization, or the government.

In addition to a familiarity with and access to mobile phones, there remain important gaps in nutrition among the children in the sample—the mean child height-for-age z-score (HAZ) is -1.4 and 29.6% of measured children are stunted (HAZ below -2). Consistent with these deficits in child nutrition, the data also highlight important areas of non-compliance with infant and young child feeding (IYCF) guidelines: only 1.3% of children are exclusively breastfed for the first six months, 21.4% of children 6-23 months of age received food from the recommended four or more food groups and just 45.5% of children 6-23 months of age met the minimum meal frequency guidelines given their age and breastfeeding status during the day preceding the survey. Adult knowledge about IYCF practices is likewise incomplete; on average, females answered 55.9% and males answered 48.5% of the 11 IYCF knowledge questions included in the baseline survey correctly.

Evaluation Design

The aim of the impact evaluation is to assess the impact, cost-effectiveness and commercial viability of mNutrition. The evaluation is being conducted by a consortium of researchers from GAMOS, IDS, and IFPRI, and it relies on a variety of different tools and methods to collect evidence on the impact of the mNutrition intervention in Tanzania. Broadly, the evaluation can be classified into three distinct but closely integrated components: a qualitative component, a quantitative component, and a business model and cost-effectiveness component.

This report focuses on the quantitative component which employs a cluster randomised controlled trial to identify the causal effect of the programme on nutrition knowledge, Infant and Young Child Feeding (IYCF) practices, women's dietary diversity, and the nutritional status of young children. Surveyed households in villages randomly assigned to the treatment group are offered access to the mNutrition content on a mobile phone, free of charge, through a door-to-door, in person visit; households in villages randomly assigned to the control group do not receive any offer of access to the programme.

In addition to the village level randomization, the evaluation also includes a second stage household level randomization among eligible households in treatment villages to help us understand how nutrition information flows between spouses. For households in treatment villages where both the pregnant woman (or the mother of the child under twelve months of age) and her spouse own distinct mobile phones, we randomly assign half to receive the mNutrition programme on just the phone of the female household member and the other half to receive the mNutrition programme on both the phone of the female household member and her spouse. Comparing outcomes between these two sub-treatment groups will allow us to explore how nutrition information flows within households and how household members form beliefs about nutrition.

The quantitative evaluation will answer the following research questions:

1. What is the impact of the mNutrition programme on women's dietary diversity?

2. What is the impact of the mNutrition programme on Infant and Young Child Feeding (IYCF) practices?

3. What is the impact of the mNutrition programme on nutritional status for children under twelve months of age at baseline?

4. What is the impact of the mNutrition programme on nutrition knowledge, including knowledge of IYCF practices, among pregnant women and the caregivers of young children?

5. Does sending the mNutrition programme to the mobile phones of both the primary female—either the pregnant woman or the mother of the child under twelve months of age at baseline—and the primary male—typically the spouse of the primary female—have a differential impact on the other primary and secondary outcomes?

Baseline Data Collection

The baseline data collection included two separate exercises: the community listing exercise (CLE) and the household survey. The survey firm OPMT led the CLE and household survey preparations and fieldwork in cooperation with the quantitative evaluation team from IFPRI.

The CLE data collection team interviewed 23,592 households in 180 different villages from the 3 rural districts in the Iringa region Tanzania between October 3rd, 2016 and November 18th, 2016. Of these households, 4,260 (18%) were identified as being eligible to participate in the study. The inclusion criteria were that the household must: (1) have a female household member who was currently pregnant or the mother of a child under twelve months of age; (2) own a mobile phone; and (3) have at least one household member literate in Swahili. From the full list of eligible households, the household survey was conducted on a random sample. In total, the baseline survey was completed in 2,833 households, 948 with a pregnant woman and 1,885 with the mother of a child under twelve months of age.

The offer of access to the mNutrition programme was made to surveyed households in treatment villages upon completion of the baseline household survey. The primary female respondent was given a brief description of the mNutrition content and asked whether she was willing to give her consent for the household to receive the messages on a mobile phone. She was also informed that the messages would be free of charge to the household. The description of mNutrition and the request for consent to receive the programme was always done at the end of the household survey. No questions were asked after introducing mNutrition to the respondent. Of the 2,141 interviewed adults in treatment villages who were asked for their consent to receive the mNutrition intervention on a mobile phone, 2,135 (99.86%) agreed to be registered for the programme.

Sample Characteristics and Implications for mNutrition

This baseline report introduces the evaluation context, describes the mNutrition intervention in Tanzania, details the evaluation design plans, summarizes the data obtained through the baseline household survey, and tests whether the randomizations successfully balanced baseline observable characteristics across the different treatment arms. The findings from this quantitative baseline report will be integrated with those from the baseline qualitative report and the baseline business model and cost-effectiveness reports to produce a mixed methods baseline report that reflects the full mNutrition impact evaluation in Tanzania. This section highlights some of the key findings from the quantitative baseline data collection.

Key Highlights:

- Sample households are resource constrained: 53.1% of households are below 150% of the national poverty line and 68.3% live on less than $2.00 per day in 2005 US dollars.
- Undernutrition is prevalent among measured children. The mean height-for-age z-score is -1.4 and 29.6% are stunted (height-for-age z-score below -2).
- Compliance with infant and young child feeding practice guidelines varies across indicators. Just 1.3% of children 0-5 months of age are exclusively breastfed and only 21.4% of children between the ages of 6-23 months received food from four or more food groups during the day preceding the survey, but 76.4% of children under five were breastfed within one hour of birth.
- Nutrition knowledge for both men and women is low, suggesting there is scope for the provision of information to improve knowledge and beliefs.
- Access to and familiarity with mobile phones is high, but substantially higher for men than women. 58.9% of female respondents and 90.7% of male respondents own a mobile

phone, while 62.4% of females and 83.5% of males received an SMS message during the two weeks preceding the survey.

82.0% of female respondents report that they would trust breastfeeding, complimentary feeding, or general health information that they received from automated SMS messages sent by an NGO, private organization, religious or voluntary organization, or the government, emphasizing the potential for the mNutrition content to affect beliefs and behaviours.

- **Demographics, assets, and wealth:** On average, households in the sample have 5.3 members, household heads are 38.6 years of age, and the primary female respondent (the pregnant woman or the mother of the child under twelve months of age) is 27.0 years of age. 74.2% of household heads are in monogamous marriages and another 13.9% are in polygamous marriages; similarly, 67.7% of primary female respondents are in monogamous marriages and 8.9% are in polygamous marriages. Crop production is the most common principal activity for household members: 73.1% of household heads and 76.8% of the primary female respondents list crop production as their primary activity. To help summarize asset ownership and wealth, we use the Progress out of Poverty Index (PPI), a tool designed to calculate the probability that a household is living below different national and international poverty lines based on their responses to ten survey questions. PPI scores suggest that 53.1% of sample households are below 150% of the national poverty line in Tanzania and 68.3% are living on less than $2.00 per day in 2005 US dollars. The sample is therefore quite resource constrained.

- **Access and use of mobile phones:** 58.9% of the primary female respondents and 90.7% of primary male respondents (typically the spouse of the primary female) report owning a mobile phone. The male respondents are substantially more likely to report having used a mobile phone to make (89.6% for males; 70.9% for females) and receive phone calls (92.7% for males; 76.5% for females) or send (64.5% for males; 48.6% for females) and receive (83.5% for males; 62.4% for females) text messages than their female counterparts over the fourteen days preceding the interview. Despite the large gap in mobile phone ownership and use between primary male and primary female respondents, females are more likely, though the difference is not statistically significant, to have used a mobile phone to receive health advice during the preceding fourteen days (8.5% for primary females; 8.0% for primary males). This suggests that while it may be easier to reach male household members in our sample through a mobile phone, the provision of nutrition information through a mobile phone may not be more effective when sent to males than when sent to females.

- **Child Anthropometry:** Measured children were 5.5 months of age, on average, and 48.8% were male. 29.6% of the children that were measured are stunted (Height-for-age Z-score below -2); according to the WHO, this is between a medium and high prevalence of stunting for a population (WHO, 1997). 12.1% of the children are underweight (Weight-for-age Z-score below -2)—a medium level of underweight prevalence. Wasting (Weight-for-height Z-score below -2) is extremely rare in the sample: just 2.0% of measured children are wasted, a low level of wasting prevalence. This is likely driven by the low heights among measured children. Together, the child anthropometry data indicates that undernutrition is an important problem for the children in the sample.

- **Women's Dietary Diversity:** Women's dietary diversity scores (WDDS), which are intended to reflect the nutritional quality of a woman's diet by measuring the likelihood that the woman is achieving micronutrient adequacy, are quite low among surveyed females. The mean number

of food groups (out of nine) was 3.5, on average, suggesting that most women in the sample would not achieve the minimum level of women's dietary diversity (MDD-W), which requires women to consume food from at least five (out of ten) food groups, for their diet to be designated as adequate (FAO, 2016).

- **IYCF Practices:** We measure IYCF using maternal reports for current and past practices for up to two children of the primary female respondent under the age of five. There is considerable heterogeneity across the six IYCF practices we consider: only 1.3% of children were exclusively breastfed for the first six months, 21.4% of children between the ages of 6-23 months receive food from four or more food groups, 45.3% of children satisfy minimum meal frequency for their age and breastfeeding status, 69.6% of children 6-8 months of age received solid, semi-solid, or soft foods, 76.4% of children were put to the breast within one hour of birth, and 83.3% of children between the ages of 12-15 months still receive some breastmilk. As a result, we expect there to be more scope for information on some IYCF practices—for example, exclusive breastfeeding and children's dietary diversity—to change behaviour.

- **Nutrition Knowledge:** We convert the eleven survey questions related to nutrition and IYCF knowledge and beliefs into a summary measure of overall nutrition knowledge for each respondent: the percent of answers that were correct. On average, female respondents answer just over half of the questions correctly (56.0%); male respondents perform slightly worse (48.5% correct), but there is room to improve nutrition knowledge for adults of both men and women in the sample.

Baseline Balance in Observable Characteristics

A successful randomization should ensure balance across the treatment arms in observable and unobservable characteristics at baseline. We assess balance in observable characteristics at baseline using two measures: the p-value from a test of the null hypothesis of no difference in means between treatment and control households and the normalized difference—the ratio of the difference in means to the average of the within treatment group standard deviation. We follow the economics literature and consider p-values above 0.05 as being indicative of balance for the characteristic being tested; we also use the rule-of-thumb proposed in Imbens (2015) and deem normalized differences below 0.25 in absolute value as evidence of balance across the treatment groups.

Village Level Randomization

The village level randomization successfully balanced observable characteristics between the treatment and control groups according to both of the two measures discussed above. Of the 204 characteristics tested across the village level treatment groups, we reject the null hypothesis of no difference in means at the 5% level for just 12. This is a rejection rate of 5.8%, or almost exactly what we should expect to uncover by chance. Normalized differences, which are sample size and scale invariant, indicate a similar level of balance: just two of the 204 normalized differences are greater in absolute value than 0.25.

Household Level Randomization

The household level randomization, conducted just within eligible treatment households, also suggests there is good balance in observable characteristics between treatment households randomly assigned to receive the mNutrition programme on just the mobile phone of the primary

female and households assigned to receive the mNutrition programme on both the mobile phone of the primary female and the mobile phone of her spouse. Of the 204 variables tested, we reject the null hypothesis of no difference in means between the two sub-treatment groups for 9 at the 5% level. This rejection rate (4.4%) is slightly smaller than the 5% we should expect by chance. Similarly, just 9 of the 204 variables have normalized differences above the 0.25 threshold.

Table of contents

List of tables, figures, and boxes

List of abbreviations

ANCOVA	Analysis of Covariance
ART	Antiretroviral Therapy
BCG	Bacillus Calmette-Guerin
CAPI	Computer Assisted Personal Interviewing
CCBRT	Comprehensive Community Based Rehabilitation in Tanzania
CDC	US Centers for Disease Control and Prevention
CLE	Community Listing Exercise
COSTECH	Commission for Science and Technology
COUNSENUTH	Centre for Counselling, Nutrition and Health Care
cRCT	Cluster Randomized Controlled Trial
DFID	Department for International Development
DHS	Demographic and Health Survey
DPT-HB	Diphtheria, Pertussis, and Tetanus
EGPAF	Elizabeth Glaser Pediatric AIDS Foundation
FANC	Focused Antenatal Care
GAIN	Global Alliance for Improved Nutrition
GSMA	Groupe Speciale Mobile Association
HAZ	Height-for-age Z-score
HNI	Human Network International
HPHB	Healthy Pregnancy, Healthy Baby
ICT	Information Communication Technology
IDS	Institute of Development Studies
IFPRI	International Food Policy Research Institute
IRB	Institutional Review Board
ITT	Intent to Treat
IUD	Intrauterine Device
IVR	Interactive Voice Response
IYCF	Infant and Young Child Feeding

JHCCP	Johns Hopkins Center for Communication Programs
LATE	Local Average Treatment Effect
mHealth Tanzania-PPP	mHealth Tanzania Public-Private Partnership
MMR	Measles, Mumps, and Rubella
MNO	Mobile Network Operator
MoHCDGEC	Tanzanian Ministry of Health, Community Development, Gender, Elderly and Children
NGO	Non-governmental Organization
OLS	Ordinary Least Squares
OPM	Oxford Policy Management
OPMT	Oxford Policy Management Tanzania
PPI	Progress out of Poverty Index
RCT	Randomized Controlled Trial
SMS	Short Message Service
TCCP	Tanzania Capacity Communication Project
TFNC	Tanzania Food and Nutrition Centre
TOR	Terms of Reference
UNAIDS	United Nations AIDS
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WAZ	Weight-for-age Z-score
WDDS	Women's Dietary Diversity Score
WHO	World Health Organization
WHZ	Weight-for-height Z-score
WN	Wazazi Nipendeni
2SLS	Two Stage Least Squares

1 Introduction

A mounting body of evidence links early childhood undernutrition to increased morbidity and mortality (Pelletier et al., 1995) as well as to poor adult outcomes including shorter stature, decreased educational attainment, reduced economic productivity (Alderman, 2006; Victoria et al., 2008; Hoddinott et al., 2013), increased incidence of non-communicable disease (Barker et al., 1989; Gluckman and Hanson, 2004). Despite the potentially serious consequences, early childhood malnutrition remains common around the world: as of 2011, 165 million children under the age of five were stunted and 52 million children under the age of five were wasted (Black et al., 2013).

Though the causes are inarguably complex, poor maternal nutrition during pregnancy (Black et al., 2013; Christian et al., 2013) and inadequate Infant and Young Child Feeding (IYCF) practices (Bhutta et al., 2013) are thought to be two of the principal drivers of early childhood undernutrition. Improving these behaviours therefore seems likely to generate important returns for both childhood nutrition and adult well-being.

With the rapid increase in access to and ownership of mobile phones across sub-Saharan Africa and the broader developing world (Pew Research Center, 2015), Information Communication Technology (ICT) interventions using mobile phones are increasingly seen as a feasible way to disperse information to individuals and households. Largely, though not exclusively, these campaigns have focused on improving farmers' information for agriculture through the provision of crop and input prices, weather information, and agricultural extension services (Svensson and Yanagizawa, 2009; Fafchamps and Minten, 2012; Courtois and Subervie, 2015; Hildebrant et al., 2015; Aker et al., 2016; Cole and Fernando, 2016).

Though less common, ICTs and specifically SMS-based information interventions have also been used to provide health-related information (Labrique et al., 2013). Typically, these interventions target improved patient drug adherence (Nglazi et al., 2013) or behaviour change related to sexual and reproductive health (Rokicki et al., 2017). Few SMS-based message campaigns have targeted nutrition-related behaviour change. Jiang et al. (2013) and Flax et al. (2014) test whether two such interventions influence Infant and IYCF practices in Nigeria and China, respectively. To date, the existing research on ICTs for nutrition and health finds mixed results on their effectiveness and the nutrition-focused ICTs have not been designed to test for impacts on child nutrition outcomes.

mNutrition, a global initiative supported by DFID, organized by GSMA, and implemented by in-country service management organisations in cooperation with mobile network operators (MNOs), explores the potential to use mobile technology to change attitudes, knowledge, behaviours, and practices for improved nutritional status. In Tanzania, the programme focuses on the provision of nutrition and health information and services to vulnerable pregnant women and caregivers of children under the age of five on their mobile phones with the goal of improving nutrition outcomes and behaviours for mothers and young children.

1.1 Objectives

The mNutrition evaluation is intended to understand and measure the impact, cost-effectiveness, and commercial viability of the mNutrition product using a mixed-methods evaluation design. The evaluation includes a quantitative component, a qualitative component, and a business model analysis. The evaluations are being conducted by a consortium of researchers from GAMOS, the Institute of Development Studies (IDS), and the International Food Policy Research Institute (IFPRI). The team draws on a number of methods and interlinked work streams to gather evidence about the impact of the mNutrition intervention in Tanzania.

- The **quantitative impact evaluation**, employing a cluster randomized controlled trial to determine the causal effect of the programme.

- A **qualitative impact evaluation**, which consists of three qualitative data collection rounds (i.e., an initial exploratory qualitative study, in-depth case studies at midline, and rapid explanatory qualitative work after the quantitative endline survey data collection) and aims to provide understanding of the context, underlying mechanisms of change and the implementation process of mNutrition.
- A **business model and cost-effectiveness evaluation** employing stakeholder interviews, commercial and end-user data, document analysis and evidence from the quantitative and qualitative evaluations to generate a business model framework and estimate the wider imputed benefits from the value-added service for the range of stakeholders involved.

The quantitative component of the evaluation is designed to contribute evidence to help answer the first two broad questions specified in the Terms of Reference (TOR, Annex G):

1. What are the impacts and cost-effectiveness of mobile phone-based nutrition and agriculture services on nutrition, health and livelihood outcomes, especially among women, children and the extreme poor?

2. How effective are mobile phone-based services in reaching, increasing the knowledge, and changing the behaviour of the specific target groups?

The quantitative component of the evaluation is best suited to address these overall study questions. The quantitative evaluation's approach to generating rigorous impact estimates of the program's effects using a causal research design will directly respond to question 1. These impact estimates will be combined with information on the cost of implementing the program components to construct estimates of cost effectiveness. The surveys from the quantitative evaluation will also provide the information needed to answer question 2 on the effectiveness of the service at reaching target groups and changing their knowledge and behaviours around nutrition.

In Tanzania, where the research consortium is evaluating mNutrition within a broader mHealth programme, the intervention aims to promote behaviour change around maternal and early childhood health and nutrition. The target group is therefore comprised of pregnant women and caregivers of children under the age of five years who reside in rural areas of the study region (Iringa).

For the sample of pregnant women and caregivers of young children in rural Iringa that are selected to participate in the quantitative study, the evaluation focuses on estimating the causal effect of access to the mNutrition programme. That is, the evaluation will identify how nutrition related behaviours, knowledge, and outcomes are altered for programme beneficiary households relative to their counterfactual levels: what the value of the outcome would have been for beneficiary households in the absence of access to the mNutrition programme.

The intended audience for the quantitative baseline report is DFID, along with other organizations involved in mNutrition and mHealth programmes globally (including local MNOs and NGOs implementing mNutrition services), national governments—in particular, the Tanzanian Ministry of Health, Community Development, Gender, Elderly and Children and the Tanzania Food and Nutrition Centre in Tanzania—international agencies and donors, and community-level health workers.

1.2 Research Questions

While the quantitative evaluation is designed to produce evidence to contribute to the broader research consortium's answers to the first two questions listed in the TOR, IFPRI also specified a set of primary and secondary research questions that will be answered using information collected by the quantitative research team. For each of the primary and secondary research questions, the evaluation focuses on estimating the causal impact of the offer of access to the mNutrition programme and of registration for the mNutrition content among households induced to participate in the programme by the treatment offer. The primary research questions that will be addressed through the quantitative evaluation are:

1. What is the impact of the mNutrition programme on women's dietary diversity?

2. What is the impact of the mNutrition programme on Infant and Young Child Feeding (IYCF) practices?

3. What is the impact of the mNutrition programme on nutritional status for children under twelve months of age at baseline?

In addition, the evaluation will also address the following secondary research questions:

4. What is the impact of the mNutrition programme on nutrition knowledge, including knowledge of IYCF practices, among pregnant women and the caregivers of young children?

5. Does sending the mNutrition programme to the mobile phones of both the primary female—either the pregnant woman or the mother of the child under twelve months of age at baseline—and the primary male—typically the spouse of the primary female—have a differential impact on the other primary and secondary outcomes?

The three primary research questions specify the main outcomes that will be studied under the quantitative component of the evaluation. These directly contribute to answering the first overall study question of the mNutrition evaluation (see Objectives, above). The first secondary research question (question 4) addresses overall study question 2 on the effectiveness of the interventions for improving nutrition knowledge and behaviours. The next secondary research question (question 5) addresses both of the overall study questions by examining how changing the target recipient of information in the household affects the impact and cost-effectiveness of the program across each of the study outcomes.

The remainder of this document proceeds as follows. Section 2 discusses the mNutrition programme being evaluated and provides basic information about the study context. Section 3 presents the evaluation design and empirical strategy. Section 4 summarizes the baseline data collection. Section 5 presents summary statistics and tests for balance across the village-level treatment groups for the outcomes and other key indicators. Section 6 explores observable differences between the overall study sample and the sample eligible for a household-level sub-randomization and tests for balance between the two treatment groups generated by the household-level randomization, and Section 7 concludes.

2 The mNutrition Intervention in Tanzania

mNutrition is a global initiative supported by DFID, organized by the Groupe Speciale Mobile Association (GSMA), and implemented by in-country mobile network operators (MNOs) and third party organisations to use mobile technology to improve the health and nutritional status of children and adults in low-income countries around the world. mNutrition is implemented through mAgri and mHealth programmes in several countries throughout sub-Saharan Africa and South Asia. The nutrition content aims to promote behaviour change around key farming practices and around dietary and child feeding practices that are likely to result in improved nutritional health within a household.

In Tanzania, mNutrition is implemented through the `Healthy Pregnancy, Healthy Baby' (HPHB) SMS text messaging service. The mass media programme accompanying the service is called Wazazi Nipendeni. The Wazazi Nipendeni programme is a US Centers for Disease Control and Prevention (CDC) funded project bringing together multiple partners contributing towards shared goals. Phase 1 of the programme, launched in 2012, was initially developed in coordination with the Tanzania Capacity Communication Project (TCCP), a United States Agency for International Development (USAID) funded programme led by the Johns Hopkins Center for Communication Programs (JHCCP). Wazazi Nipendeni was one of several behaviour change programmes using methods as diverse as TV drama series, radio distance learning for community health volunteers and several integrated mass media campaigns. The mass media campaign was developed by JHCCP, while the SMS component of the campaign was led by the mHealth Tanzania Public-Private Partnership (mHealth Tanzania-PPP). The public-private partnership was initiated by the Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC), with financial support from CDC.[1] Wazazi Nipendeni is available nationally and on all phone networks.

The HPHB SMS service sends free text messages with health care information to pregnant women, mothers with newborns, and male supporters and general information seekers in Tanzania to drive health-seeking behaviour (Open Government Partnership, n.d.). The SMS messages are sent in Swahili, originally to women up to 16 weeks post-partum on a range of pregnancy and early childhood issues times to the stage of the pregnancy and the age of the child. Anyone interested in receiving healthy pregnancy information can text the word `MTOTO' (child) to the short code 15001. Registrants receive instructional messages, allowing them to indicate the woman's current week or month of pregnancy (or the age of the newborn baby) during the enrolment process. This process allows the recipients to receive specific text messages relevant to the time and stage of pregnancy or early childhood. The message frequency also varies depending on the life stage of the woman and child, ranging from nearly daily during key periods of pregnancy to less than weekly for mothers of children over the age of two.

The mNutrition programme has supported mHealth projects in 8 countries through the development of nutrition content, and GSMA has assisted projects with product development primarily through user experience research and business intelligence support. Nutrition related content was a small component of the original HPHB SMS service but was expanded substantially with the addition of the content contributed through GSMA under the mNutrition programme. mNutrition adds roughly 120 nutrition messages which are delivered to mothers or caregivers of children up to five years old. HPHB and mNutrition is available to households in all regions of Tanzania, on all mobile phone networks, and participating individuals receive the text messages free of charge. The resulting product will be referred to as mNutrition in the remaining sections of this report.

[1] The Wazazi Nipendeni campaign and text messaging service is funded by the US President's Malaria Initiative and US President's Emergency Plan for AIDS Relief (PEPFAR) and implemented through US Government agencies USAID and CDC. It is run in coordination with the National Malaria Control Program, National AIDS Control Program and Health Promotion and Education Section. On the ground, health facility orientation support is also provided by the US Government, Aga Khan Health Services and Canadian International Development Agency. Other implementing partners include Jhpiego, the Elizabeth Glaser Pediatric AIDS Foundation (EGPAF), the Mwanzo Bora Program, Comprehensive Community Based Rehabilitation in Tanzania (CCBRT), Tunajali Project, PLAN International, Aga Khan Foundation and others.

The 120 nutrition messages included in the mNutrition programme are drawn from 42 factsheets on nutrition-related behaviours identified as key determinants of outcomes that were developed by GAIN together with local content providers the Centre for Counselling, Nutrition and Health Care (COUNSENUTH) and Every1mobile. The information contained in these factsheets was then adapted to the context of Tanzania and made mobile ready by the local content providers under the guidance of the MoHCDGEC and the Tanzania Food and Nutrition Centre (TFNC). As a part of the adaptation process the message content was tested with potential users in Tanzania, after which the language and substance was adjusted and messages that were identified as not being useful were removed from consideration for the final programme. The message testing process highlighted the importance of replacing technical terminology that was likely to be unfamiliar to the message recipients with language that was more commonly used but that still conveyed the evidence-based content of the original factsheets.[2] Included in the final programme are messages that encourage the consumption of iron folic acid tablets during pregnancy[3] and messages that promote vitamin A rich complementary feeding practices[4] and the inclusion of animal source foods in young children's diets,[5] as well as messages providing information on other behaviours accepted as being critical determinants of nutrition outcomes.

The original HPHB text messaging service did not have the capability to deal with voice messages, but voice messages were developed as part of the local content development process in Tanzania. Under a separate agreement, GSMA commissioned Human Network International (HNI) to incorporate the mNutrition content into their 321 service, provided in partnership with Vodafone. In contrast to HPHB, the 321 service is a `pull' service, whereby users dial a shortcode and navigate through interactive menus to find the information they are seeking. The system mostly plays audio clips to users, rather than sending SMS text messages. 30 interactive voice response (IVR) scripts were selected to be integrated into the 321 health service, and were being recorded at the time of the baseline field visits. As these IVR messages are provided by HNI, they are not included in the SMS-based programme that this study evaluates.

While the mNutrition intervention has been available in Tanzania since early 2017, large-scale marketing for the programme has not yet been completed and discussions with the in-country implementing organisations indicated that take-up in the study region (Iringa) was likely to be low. Data collected through our baseline survey—discussed in more detail below—confirm the low levels of take-up in rural Iringa: just 7.2% of female respondents and 6.7% of male respondents report having received any SMS-based nutrition messages during the previous two years. As the combined HPHB+mNutrition programme is not the only available SMS-based health and nutrition content, these figures almost certainly overstate the baseline exposure to the intervention in the study region.[6]

2.1 Study Context

The United Republic of Tanzania is an East African nation with an estimated population of 55.6 million (2016), just under 70% of which reside in rural areas (World Bank, 2017). As of 2016, 64% of working age males and 70% of working age females were employed in agriculture (ILO, 2017), with the main agricultural export commodities including coffee, tobacco, tea, cotton, and sisal (FAOSTAT, 2014). Tanzania is divided into 31 regions and regions are further subdivided into a total of 169 districts.

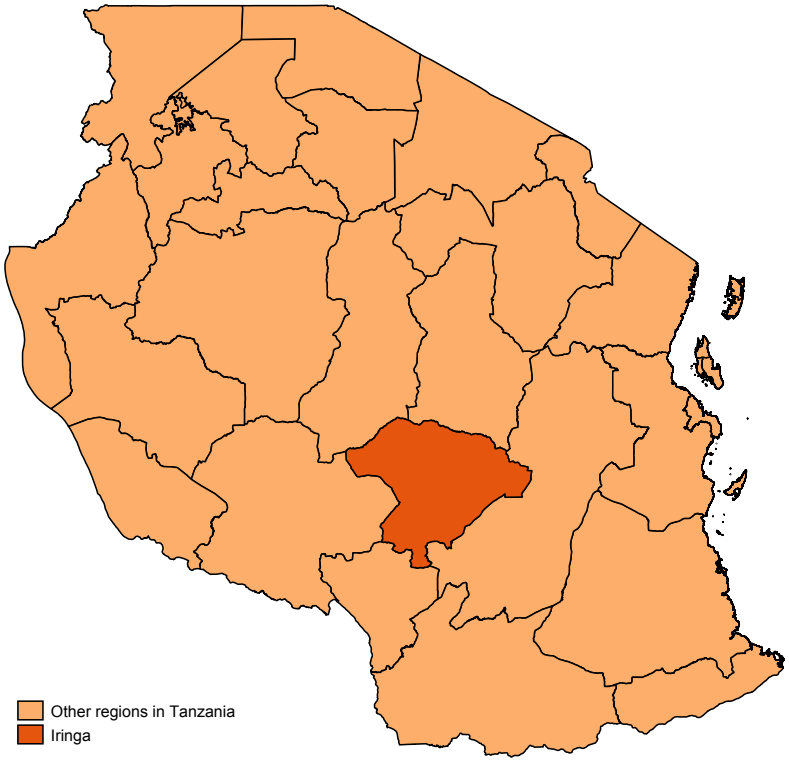
[2] For example, messages discussing Anemia use the term "low blood" and diarrhea is referred to as "running stomach."
[3] "Mom, start using iron folic acid tablets within the first 6 weeks of pregnancy to reduce risks of brain or spinal disorders to your baby."
[4] "Dear mother, remember to give your baby fruits and vegetables containing Vitamin A such as pumpkin, carrots, spinach, matembele, papaya or mango."
[5] "Build the practice of giving your baby natural animal foods like meat, fish, poultry, seafood, dairy products and eggs because it is important for the growth of their body and mind."
[6] Admittedly, there is likely to be some overlap in content between the mNutrition programme and any other SMS-based nutrition information programmes operating in Iringa region. We still view 7.2% (among females) and 6.7% (among males) as being acceptably low levels of exposure to existing SMS-based nutrition programmes, even if the exposure is to the mNutrition content itself or an identical alternative programme.

Child undernutrition is a pervasive problem in Tanzania, particularly among young children. In the 2016 Demographic and Health Survey (DHS), 34.4% of children under five were identified as being stunted (height-for-age z-score below -2). Wasting is less common, with only 4.4% of measured children under five having a weight-for-height z-score below the -2 threshold. Over half of measured children 6-59 months of age (57.7%) are anaemic (DHS, 2016).

This study takes place within the three rural districts of the Iringa region in Tanzania: Iringa rural, Kilolo, and Mufindi. Figure 1 displays the location of Iringa region in Tanzania. Iringa became an independent region in 1964, before which it was a part of the Southern Highland Province. As of the 2012 Population and Housing Census, the total population of Iringa region was estimated to be 0.9 million, 73% of which resided in a rural area. Agriculture is the primary means of livelihood for most households in the rural parts of the Iringa region: roughly 89 percent of households in Iringa rural, 92 percent of households in Mufindi, and 92 percent of households in Kilolo are involved in agriculture. Average household size was 4.2 and the adult literacy rate was 79% among the rural population (PHC, 2012).

As in Tanzania more broadly, child undernutrition is a severe problem in Iringa. 41.6% of children under the age of five in the 2016 DHS were stunted. This figure is nearly 7 percentage points higher than the national average, suggesting child malnutrition may be more prevalent in Iringa than elsewhere in Tanzania. Additionally, 3.6 percent of children under five were wasted and 40.3% of children 6-59 months of age were anaemic.

Figure 1: Iringa Region in Tanzania



2.2 Nutrition Programmes in the Study Region

In part, Iringa was selected as the location for the current study because of a dearth of existing relationships between mHealth Tanzania-PPP and organisations with a presence in the region. Consistent with this, the mNutrition baseline qualitative report finds that households in study villages typically rely on health workers at local health clinics and community health workers for their nutrition information needs. Often the information from health clinic workers is received during antenatal visits, which also involve the

provision of non-nutrition related information, testing, and other services, sometimes leaving little time for issues related to nutrition (Barnett et al. 2017).

Though at the start of the study there was limited availability of nutrition information and nutrition programmes, the government of Tanzania has prioritized improving nutrition outcomes nationwide through different initiatives. For example, in 2011 Tanzania became a part of Scaling Up Nutrition, a global push to coordinate stakeholders and resources in order to improve nutrition outcomes. Partly as a consequence, in April of 2013, CONCERN Worldwide and the United Nations Children's Fund (UNICEF) began a five-year programme called "Bringing Nutrition Access to Scale in Iringa, Njombe and Mbeya" to improve child and maternal nutrition, with an emphasis on interventions that focus on the first 1,000 days of life. The planned or implemented interventions include a community-led total sanitation programme, a national simplified birth registration campaign, a radio serial drama designed to promote healthy behaviours and the use of reproductive, maternal, new-born and child health services, and the provision of social and behaviour change communication about infant and young child feeding practices at the community and health facility level (Concern 2015; UNICEF 2016).

3 Evaluation Design

3.1 Overview of the Quantitative Evaluation Design

The quantitative evaluation is designed as a cluster randomized controlled trial, with two stages of randomization: a village level randomization where villages are assigned to a treatment group or to a control group and a household level randomization within treatment villages whereby households are either assigned to receive the mNutrition content on just the mobile phone of the primary female or on the mobile phones of both the primary female and the primary male. In villages that were assigned to the treatment group, sampled households are offered access to the mNutrition content on their mobile phone, free of charge, through a door-to-door, in person visit. In villages that were assigned to the control group, no offer of access to the programme is made. Though registration for the mNutrition programme is possible for all households—regardless of treatment assignment—pre-baseline discussions with the organization implementing the mNutrition programme in Tanzania suggested that take-up of their existing programme was low in the study region. We therefore believe that the random assignment of access to the programme is likely to generate a substantial difference in participation between treatment and control households.

Randomized controlled trials (RCTs) are widely viewed as the most reliable method of quantitative programme evaluation. By allowing researchers to randomly select who is affected by or offered access to a programme, RCTs provide a clean way to estimate the causal effect of the programme. In comparison, observational studies are forced to compare outcomes across individuals or groups that *choose* to participate in the programme to individuals or groups that do not. In most circumstances, this simple difference will be the sum of the causal effect of participating in the programme and a term representing how the individuals or groups that choose to participate are different from those that do not, even in the absence of exposure to the programme. This latter term, typically referred to as selection bias, can be positive or negative depending on the context and the intervention being studied. Selection bias is eliminated by the random assignment of treatment (or a treatment offer) in an RCT; the individuals or groups included in the study are allocated to receive the treatment (or treatment offer) by the randomization mechanism controlled by the research team. On average, the units assigned to the treatment group and the units assigned to the control group should therefore have the same outcome levels, in the absence of exposure to the treatment. As a result, any difference in outcomes at the end of an RCT can be attributed to the intervention being evaluated and not to selection bias.

Prior to fieldwork, sample size calculations were undertaken to identify the number of study clusters to include in the evaluation. To facilitate the identification and sampling of potential beneficiary households in both treatment and control villages, baseline data collection included two separate activities. First, a community listing exercise (CLE) was conducted to gather basic information to determine whether households were eligible to be sampled. Next, a household survey collected detailed information on all outcomes as well as on other characteristics likely to be predictive of outcomes for sampled households. Between the CLE and household survey, the research team processed the CLE data to identify eligible households, randomly sampled households from those deemed eligible, and conducted the village level randomization in such a way as to minimize the expected variance of the treatment effects. As data from the household survey became available, households in villages assigned to the treatment group that met an additional eligibility requirement were randomly assigned to one of two different household level treatment arms. Below, we provide more information about each step in the quantitative evaluation's basic design and baseline fieldwork.

3.2 A Cluster Randomized Controlled Trial

To identify the impact of the mNutrition programme on nutrition related behaviours, knowledge, and outcomes, the quantitative evaluation includes two rounds of household surveys: a baseline survey which was completed between October 2016 and December 2016 and an endline survey with the same set of

households to be completed between October 2018 and December 2018. We elected to schedule the baseline data collection (and the start of the intervention) two years before the endline data collection to ensure that beneficiary households would be exposed to the mNutrition programme for sufficient time to potentially affect households' beliefs, behaviours, and key nutrition outcomes that are likely to be affected by the beliefs and behaviours of household members. Generating impacts on child anthropometry, in particular, is likely to require sustained behaviour change throughout the first 1,000 days of children's lives (de Onis et al., 2013). Each round of data collection will collect information on all primary and secondary outcomes, as well as on individual and household characteristics that are likely to be strong predictors of the primary and secondary outcomes. While each round of data collection serves a somewhat distinct purpose, collectively they produce the information necessary for a systematic and statistically well-powered investigation of the causal impact of the mNutrition programme on outcomes. Below, we list the two rounds of data collection along with the key motivations for each.

Baseline Household Survey (October 2016-December 2016):

- Assess balance in observable characteristics across treatment groups in the cRCT study
- Use baseline measures of outcomes in endline analyses to improve statistical power through ANCOVA or difference-in-differences specifications
- Use baseline measures of non-outcome characteristics to reduce residual variance and potentially improve precision of the treatment effect estimates
- Explore the relevance of the mNutrition programme for the study population by identifying gaps in nutrition knowledge and practices, measuring mobile phone access, assessing households' interest and hypothetical trust in nutrition information received through SMS messages

Endline Household Survey (October 2018-December 2018):

- Measure primary and secondary outcomes to enable estimation of the programme's impacts
- Connect treatment effects to levels and changes in key individual and household characteristics to better understand causal pathways

To ensure that the evaluation accurately measures the causal impact of access to the mNutrition programme, the quantitative evaluation is based on a cluster randomized controlled trial (cRCT). From the randomly selected sample of villages participating in the evaluation, IFPRI randomly assigned households in half the villages to a treatment group (T)—where sampled households received a door-to-door offer of access to the content—and households in the other half of participating villages to a control group (C) that did not receive a similar offer. Because the assignment of villages was random, any average difference in outcomes between households in the two groups can be attributed to the difference in access to the mNutrition programme.

In addition to the village level randomization, the evaluation also includes a second stage household level randomization within treatment villages; households in treatment villages where both the mother of the young child or pregnant woman (the primary female) and the primary male[7] own distinct mobile phones (and were surveyed) were randomly allocated to either just receive the mNutrition content on the mobile phone of the primary female (T-F), or to receive the mNutrition content on the mobile phone of the primary female and the mobile phone of the primary male (T-F+M). By comparing behaviours and outcomes between treatment households in the T-F+M group and those in the T-F group, learning will be generated about how information flows between spouses.

[7] The primary male is typically the spouse or partner of the primary female. In cases where the primary female had no spouse/partner, or that spouse/partner was unavailable, enumerators were instructed to use either the male head of household, the father, the father-in-law, or the brother of the primary female as the primary male respondent.

The random allocation of villages to the treatment or control group and treatment households to the T-F group or the T-F+M group should ensure that a simple difference between treatment arms yields unbiased estimates of the causal effect of the programme. However, while the theoretical basis for estimating causal impacts through a randomized experiment is clear, the data collected through the baseline survey offer an opportunity to empirically assess how likely it is that the random assignment to treatment arms was successful at identifying similar groups of households. We can do so by comparing baseline measures of the primary and secondary outcomes—as well as other observable characteristics likely to be strongly correlated with the outcomes—across treatment groups. If we find that there are few differences between households in the different groups, it sends a strong signal that the evaluation is likely to be able to estimate the unbiased causal impacts of the programme.

3.3 Household Eligibility and Village Sampling

3.3.1 Household Eligibility Criteria

In Tanzania, the mNutrition messages have been integrated within the existing HPHB programme that targeted general health information to pregnant women, mothers of children up to age five, and their caregivers. Though the mNutrition content is designed to benefit children up to the age of two, we elected to restrict potential beneficiary households to those with either a pregnant woman or a child under twelve months of age at the time of the CLE.[8] This age restriction was made to ensure that all of the focus children included in the sample would have at least one full year of exposure to the programme during their first 1,000 days of life, a period during which the return to better nutrition is particularly high (Ruel et al, 2008; Black et al., 2010; Victoria et al., 2010; Ruel, 2010; Black et al., 2013).

A subset of the mNutrition content is targeted towards pregnant women—including SMS reminders to take iron supplements, consume iodized salt, and remain active during pregnancy—and other messages focus on behaviours that need to occur immediately following childbirth; for example, the programme includes messages encouraging the mother to initiate breastfeeding as soon as possible after the birth. This information is clearly only useful if received during pregnancy. We therefore decided that the evaluation would sample households in such a way as to be certain that it would be possible to identify treatment effects for behaviours and outcomes that are only relevant during pregnancy or shortly thereafter. In practice, this was done by setting two distinct sampling targets in each study cluster: one for households with pregnant women (6 per cluster) and one for households with a child under the age of twelve months (11 per cluster).

In addition to eligibility restrictions related to the age of the focus child or the pregnancy status of the primary female, several other eligibility criteria were used to identify the study sample. First, to be eligible, households were required to have at least one household member that was literate in Swahili. Second, at least one member of the household had to own a mobile phone. These restrictions were made to ensure that the mNutrition content—which is sent to mobile phones via SMS messages in Swahili—would be accessible to everyone included in the sample. In practice, these additional restrictions were rarely binding: of the 4,689 households identified as having either a pregnant woman or the mother of a child under the age of twelve months, 94.7% also owned at least one mobile phone and had a member who was literate in Swahili. All households that met the three eligibility criteria were retained in the potential study sample from which the final sample was later selected.

3.3.2 Cluster Size

The quantitative evaluation groups together all households residing within the same village by designating each sampled village as a distinct study cluster and ensuring that the treatment status is the same for all

[8] The mother of the child under twelve months of age was also required to be a household member.

households in the same cluster. The resulting cluster randomized controlled trial provides less statistical power than a randomized controlled trial that assigns treatment status at the individual level, but it greatly reduces the risk that spillovers—when one unit's treatment assignment has a direct effect on the outcomes of another unit—will contaminate estimates of the effect of the mNutrition programme.[9]

In determining the level at which to cluster observations, three potential consequences were considered. First, the mNutrition treatment is information provided through SMS messages. For households that receive the content on their mobile phones it is easy and relatively costless to share the information with other households. It was therefore not reasonable to expect that the potential beneficiaries—rural households in the Iringa region of Tanzania—would not communicate the content they received through the mNutrition programme with other households in their village. While this type of information sharing spillover across individuals within the same village is not problematic if all households in the same village are assigned to the same treatment group, it would undermine the ability of the evaluation to estimate the causal impact of the mNutrition messaging if individuals in the same village were assigned to different treatment groups. For example, without clustering together all households from the same village, a finding that there were no differences in knowledge about nutrition across households assigned to the treatment group and those assigned to control group could indicate that the mNutrition content did not increase nutrition knowledge, or it could be that treatment households received and learned from the messages, but they also shared the information with neighbours who were assigned to the control group. By ensuring that treatment allocation is constant within villages, the potential for this type of spillover is greatly reduced in a cluster randomized design.[10]

A second consideration was the need to identify a sufficient number of eligible households in each study cluster. At the beginning of baseline fieldwork, our aim was to identify six eligible households with a pregnant woman and eleven eligible households with the mother of a child under twelve months of age in each study cluster. Based on our experience piloting the baseline questionnaire in the Pwani region of Tanzania and on available data from the 2010 Demographic and Health Survey (DHS), we determined that most villages in Iringa would likely contain enough eligible households to meet these per cluster targets.

Though defining broad study clusters helps to reduce the likelihood of spillovers and increase the probability of identifying a sufficient number of eligible households in each study cluster, larger study clusters will also reduce statistical power for a given fieldwork budget. To identify the desired number of study clusters, the research team will have to include more geographic areas in the sample. This can result in potentially large increases in travel costs for data collection teams, as they are forced to travel increasing distances between study clusters. Further, if clusters are large enough, it may be impossible to identify the number desired for the evaluation. For example, Tanzania contains 31 regions, not nearly enough to satisfy the 180 study clusters that our power calculations suggest were required.[11] Ultimately, the fieldwork budget for the quantitative evaluation was sufficient to cover fieldwork costs within one region of Tanzania, and there were too few available divisions, districts, wards, and health clinics within the regions being considered for us to use any of these to define study clusters.[12]

3.3.3 Sample Size Calculations

Power calculations were conducted to determine the sample size necessary to identify meaningfully sized treatment effects of the mNutrition programme on height-for-age Z-scores (HAZ) for children. The

[9] Though clustering at the village level will unquestionably reduce the likelihood of spillovers, it is not sufficient to completely rule them out. We will assess whether information spillovers across villages are empirically relevant at endline by comparing changes in nutrition knowledge, behaviours, and outcomes between control villages that are spatially close to a greater share of villages that were assigned to the treatment group of this study and control villages where that share is smaller.
[10] Similar concerns prevented us from randomizing treatment at the sub-village level. Sub-villages within the same village share some administrative and government services and are spatially close to one another.
[11] Fewer than 180 study clusters would be needed with region level study clusters because the intracluster correlation would decline towards zero. However, even with an intracluster correlation of zero 100 study clusters would be required.
[12] Districts, divisions, and wards are the administrative divisions in Tanzania above villages. Health clinics are not an administrative division, but they generally have well defined service or catchment areas.

evaluation is designed to detect impacts on HAZ because HAZ for young children has been identified as both a good indicator of undernutrition and a meaningful predictor of later in life health and well-being (Black et al., 2013). In addition, HAZ measurement relies just on the data collection team, not on accurate reporting by a household member. Therefore, relative to self, mother, or household member reported measures of health and nutrition, HAZ is less vulnerable to problems related to measurement error or reporting bias.

Data from the 2010 Demographic and Health Surveys (DHS) (Boyle et al., 2017) in Tanzania were used to estimate critical moments of the HAZ distribution. More specifically, the control group mean and standard deviation for HAZ were estimated using the mean and standard deviation HAZ score for children 0-59 months of age in the 2010 DHS Tanzania survey. Similarly, the intracluster correlation[13] was set to 0.09 based on calculations using the 2010 DHS data from Tanzania. The autocorrelation[14] was set to 0.7 based on calculations from children 0-5 years of age, measured one year apart, from the NEEP-IE study in Malawi (Gelli, 2017) and for children measured at 6 months and again at 24 months for the PROMIS study in Mali (Huybregts, 2017); the household attrition rate between baseline and endline was assumed to be 5.0%,[15] and power calculations were conducted using the implied number of endline observations under this level of attrition. Finally, it was assumed that the cRCT would generate a take-up gap for the mNutrition programme between treatment and control villages of 70.0%, and that baseline data collection teams would be able to identify eleven eligible households with a child under twelve months of age at baseline. The resulting power calculations then indicate the required number of study clusters (villages) for a desired minimum detectable effect size[16] given the assumptions about the other parameters.[17] Ultimately, IFPRI determined that the evaluation should be able to detect differences in HAZ between the treatment and control groups of at least 0.25 standard deviations, which required that at least 180 village clusters be included in the study. An effect size of 0.25 standard deviations was selected on the basis of a previous review which suggests that several well-designed nutrition education interventions in developing countries improved child HAZ by, on average, 0.25 standard deviations (Bhutta et al., 2008); it was determined that, given the existence of these interventions, to be considered a successful way to impact child HAZ, the mNutrition programme would need to produce an HAZ effect size at least as large. We followed convention in the impact evaluation and economics literature by setting size equal to 0.05 and power equal to 0.8 (Duflo et al., 2007). Table 3.1 displays the full set of parameters as assumed or calculated while conducting power calculations.

In the second column of Table 3.1 we show the assumed or estimated (number of clusters) values for all parameters used in the power calculations. Columns 3-5 show updated power calculations using, when possible, data collected during the baseline survey.[18] For example, for all three outcomes we used the observed control group mean and standard deviation and calculated the actual intracluster correlation from the baseline data. Additionally, we use the realized sample sizes, paired with our assumed 5% attrition rate, in place of the assumed sample size shown in Column 2. For all three outcomes, the table suggests we should be able to identify a range of meaningful effects sizes. Comparing the assumed effect size for HAZ in Column 2 with the updated value based on power calculations using baseline data in Column 3, we see

[13] The intracluster correlation for HAZ is the fraction of total HAZ variance that can be explained by the within cluster (village) variance.
[14] The HAZ Autocorrelation is the correlation between child HAZ at baseline and endline, scaled by product of the standard deviation of HAZ at baseline and the standard deviation of HAZ at endline.
[15] A review of attrition in recent IFPRI randomized controlled trials in Sub-Saharan Africa identified attrition rates (over 1.5-2 years) between 4% (Gilligan et al., 2015) and 16% (Olney et al., 2016). We opted to assume that attrition would be close to the bottom of this range because of the detailed mobile phone information we were collecting through the baseline survey. By recording all phone numbers used by each household in the data, we will be able to attempt to contact households via mobile phone in the event that we are unable to locate them when the fieldwork teams visit their village.
[16] The smallest true treatment effect that a research design will be able to identify as statistically significant.
[17] The power calculations also assumed that the empirical strategy used to estimate treatment effects would be ANCOVA.
[18] The IYCF practices outcome uses a summary index over whether for the two most recent births (within the past five years) breastfeeding began within one hour of birth, whether children 6-23 months of age were fed from at least four different food groups during the day and night preceding the survey, whether children 6-23 months of age meet minimum meal frequency guidelines given their age and breastfeeding status, whether children 0-5 months are exclusively breastfed, whether children 6-8 months have been introduced to solid, semi-solid, or soft foods, and whether children 12-15 months are still being breastfed. For each indicator we demean the value by the overall mean and scale the demeaned value by the control group standard deviation. The overall IYCF index is then generated by averaging the normalized values across all six indicators.

that the study should be able to detect smaller changes in HAZ than anticipated. This improved power is driven by the smaller standard deviation and intracluster correlation for HAZ estimated from the baseline data.

Table 3.1: Parameters for Sample Size Calculations

Parameter Name	HAZ	HAZ	WDDS	IYCF Practices
	Anticipated	With Realized Parameters (when possible)		
Take up gap	0.7	0.7	0.7	0.7
Power	0.8	0.8	0.8	0.8
Size (alpha)	0.05	0.05	0.05	0.05
Number of clusters	181	180	180	180
Attrition	5%	5%	5%	5%
Total baseline observations	1895	1852	2833	2173
Total endline observations	1804	1759	2691	2064
Minimum detectable effect size	0.25	0.22	0.27	0.06
Mean control	-1.47	-1.44	3.48	-0.01
Standard deviation control	1.38	1.22	1.10	0.29
Intracluster correlation	0.09	0.07	0.11	0.02
Autocorrelation	0.7	0.7	0.1	n/a
Observations per cluster	10	9	14	11

3.3.4 Village Sampling

To identify the villages to include in the quantitative evaluation, our data collection partner Oxford Policy Management Tanzania (OPMT) provided us with a list of all rural villages in the Iringa region of Tanzania. From this list of all rural villages, we dropped three villages classified as “Miji Mdogo,” which is a term used to represent rural areas with emerging urban characteristics. We made this restriction so that the entire sample would be rural and because Miji Mdogo have a different administrative structure than rural villages, which could complicate the planned fieldwork. The remaining 354 rural villages were eligible to be selected for the study sample. Of these 354 eligible villages we randomly selected 180 to be included in the main sample and an additional 20 village to be used as potential replacements. To select villages for the sample, a uniform random number between 0 and 1 was drawn for each of the 354 potential sample villages, the villages were ranked on the basis of this random number draw, and the 180 villages with the lowest random number draws were assigned to be included in the main sample.

Ultimately, two of the villages originally selected for the main sample were replaced. In both cases, when the CLE data collection team arrived in the village they discovered that there was no mobile network coverage anywhere in the village. If these villages were assigned to the treatment group,[19] the lack of mobile network coverage would decrease the likelihood that households in the village would be able to access the mNutrition content. As a result, these villages were immediately replaced using the highest ranked replacement villages available.

[19] Treatment assignment had not yet been determined. Villages were assigned to the treatment and control groups after their CLE data had been collected.

3.4 Data Collection, Household Sampling, and Random Assignment

3.4.1 Community Listing Exercise

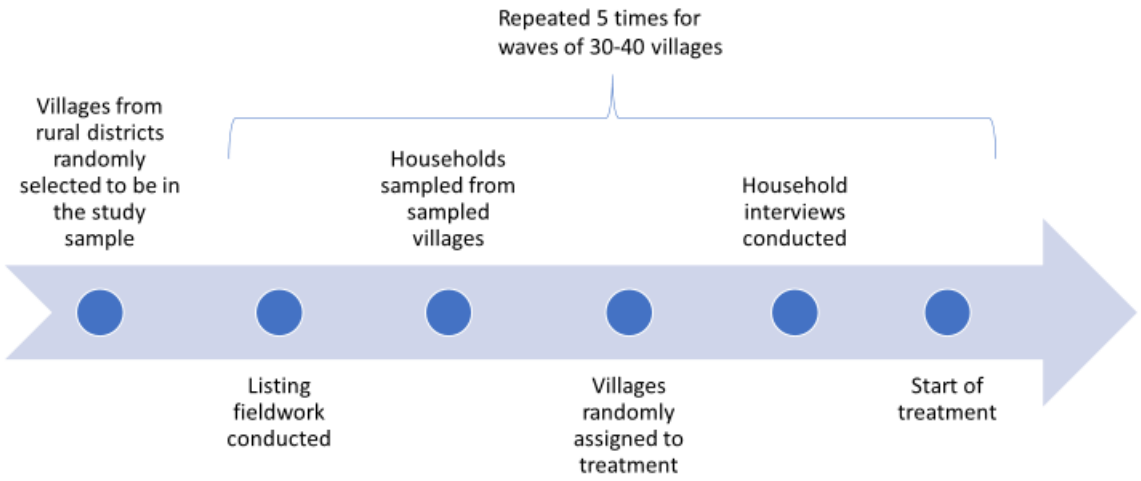
The baseline data collection included two separate exercises: the community listing exercise (CLE) and the household survey. The survey firm Oxford Policy Management Tanzania (OPMT) led the CLE and household survey preparations and fieldwork in cooperation with the quantitative evaluation team from IFPRI. OPMT has substantial experience with large scale household surveys and listing exercises in Tanzania, a thorough understanding of local contextual issues that could affect survey design and implementation, and expertise and familiarity translating technical questionnaires into Swahili.

Both the CLE and household questionnaires were designed by the IFPRI team based on past experiences conducting quantitative evaluations of nutrition interventions in Sub-Saharan Africa and taking into account the information presented in the landscape analysis (Barnett et al., 2016) conducted by IDS, the qualitative lead in this research consortium. The CLE questionnaire was devised to quickly extract the information necessary to identify whether households were eligible to participate in the evaluation as well as the geographic coordinates of the household and contact information for the household to facilitate a follow-up visit if the household was selected to be a part of the main sample.

The mNutrition programme is designed to send life stage specific nutrition information to the mobile phones of pregnant women, mothers of young children, and their caregivers. Because pregnancy and early childhood are not permanent, or even long-lasting life stages—women are only pregnant for 40 weeks and child undernutrition is thought to be the most damaging during the first two years of life (Black et al., 2013)—there was a limited period during which the intervention was likely to be relevant to potential beneficiaries. The identification of potential beneficiary households was therefore recognised as being extremely time sensitive, and care was taken to ensure as little time as possible would pass between when households and individuals in treatment villages were identified as being eligible for the mNutrition programme and when they were invited to register for the programme.

Due to the expected length of the full baseline data collection—nearly ten weeks to complete both the CLE and household survey—we determined that it would not be reasonable to wait until the full CLE had been completed to select the sample, conduct the randomization, and begin the baseline household survey; by the time the CLE was completed, many of the women who were identified as being pregnant and eligible to participate in the study would likely already have given birth. Instead, we elected to complete the baseline fieldwork on a rolling basis with two separate data collection teams: one responsible for the CLE and the other for the household survey. Typically, the CLE team passed through a village two weeks prior to the household survey team. During the time lag between the CLE and household survey, the CLE data were processed, all eligible households in a village were identified, the household sample was selected from the list of eligible households, and a new group of villages—those for which the CLE data collection had just been completed—were allocated to the treatment and control group. Figure 3.1 presents the timeline of baseline fieldwork to help illustrate how the CLE data collection was chronologically related to the other important baseline events.

Figure 2: Timeline of Baseline Fieldwork



To facilitate the completion of the baseline data collection activities the CLE was designed to collect, as quickly as possible, all of the information required to determine whether households were eligible to participate in the study. The CLE questionnaire therefore only contained the questions necessary to determine eligibility, a field to record the number for a mobile owned by the household, the birth date of the child under the age of twelve months (if the household had a member under twelve months of age), and basic preliminaries such as the GPS coordinates of the home, the interview date, the interview start and end time, and the name of the household head and the mother of the child under twelve months of age or the pregnant woman. The full text of the CLE questionnaire can be found in Annex A.

3.4.2 Household Sampling

Household sampling was done on a rolling basis as the CLE data became available.[20] From the list of households that were identified as being eligible through the CLE data collection, we separately randomly sampled households with a pregnant woman and households with a child under twelve months of age.[21] A list of potential replacement households of both types was also generated when possible for all sampled villages.

Though initially, the evaluation team hoped to sample six households with a pregnant woman and eleven households with a child under twelve months of age in each study village, it became clear early in the CLE fieldwork that it would not be possible to identify enough eligible households in each sample village; in particular, reported pregnancy rates were meaningfully lower than expected.

There were several potential causes for the lower than expected number of identified households with a pregnant woman. First, enumerators reported numerous cases of young households that had migrated to the urban capital city of Iringa region. To the extent that this rural-urban movement was concentrated among young adults who—in the absence of their urban migration—would be likely to have young children or currently be pregnant, this pattern of migration would reduce the number of pregnant women observed in

[20] There were five "waves" of sampling and randomization, the first four of which corresponded to roughly two weeks of CLE data collection and the last which included just three villages.
[21] Households that contained both a pregnant woman and the mother of a child under twelve months of age were designated as being part of the potential sample for households with a pregnant woman. This was done because of the relative shortage of eligible households with a pregnant woman.

the CLE. Second, many women in the Iringa region of Tanzania may have elected not to identify themselves as being pregnant if they had not yet reached a certain gestational age. It is a commonly held belief that it is bad luck to tell others that you are pregnant early in a pregnancy. It is likely that this underreporting of pregnancies also contributed to the low pregnancy rates observed in the CLE data.

To minimize the loss in sample size, we adjusted our sampling strategy when we were made aware of the low observed fertility in study villages. Rather than sampling exactly six households with a pregnant woman and eleven households with a child under twelve months of age, we over sampled households from villages that had more than the targeted number of eligible households of either type. In this way, we identified nearly the targeted number of total households to be interviewed through the household survey.[22]

In practice, the determination of how many households to sample was done in the following manner. If a village had at or below the targeted number of eligible households (six for households with a pregnant member and eleven for households with a child under the age of twelve months) then all eligible households of that type were selected to be a part of the sample for the household survey. If there were more eligible households than the targeted number, we sampled enough from the village such that the average number of sampled households in that wave—the group of villages where the CLE had been completed during the previous two weeks—was as close to 17 as possible. The over sampling was done by progressively increasing the number of households that are sampled from villages with extra eligible households until the average number of sampled households with a pregnant woman was six and the average number of sampled households with a child under twelve months of age was eleven among the villages in that wave.[23] In addition, up to two replacement households with a pregnant woman and up to three replacement households with a child under twelve months were identified in each village if possible.[24]

After calculating the number of households to sample in each village, households were sampled (or assigned to be replacement households) by drawing a uniform random number between zero and one for each eligible household in the village. The households were ranked on the basis of this random number draw, with separate rankings generated for households with a pregnant woman and households with a child under twelve months of age. If the rank of the household was below the village threshold for being included in the main sample then then household was assigned to be a part of the main sample.[25] If the rank of the household was higher than the number to be included in the main sample but below that number plus the number of replacement households, then the household was designated as a replacement household for that village.[26]

3.4.3 Village Level Randomization

After identifying the households to be included in the main sample and as potential replacements, villages were assigned to one of two village level treatment groups: a treatment group where all sampled households were offered access to the mNutrition programme and a control group where no offer of access to the programme was made. Though the random assignment of villages to treatment groups ensures that we will be able to estimate the causal effect of access to the mNutrition programme on outcomes, previous research suggests there may be considerable gains in precision from using more efficient treatment allocation mechanisms (Bruhn et al., 2009). Given the large variation across sample villages in the number of eligible households with a pregnant woman and the dependence of the primary outcomes on the

[22] The data collection team was not able to complete a household survey with every targeted household. The total number of completed baseline surveys was 2,833, 92.6% of the targeted number of 3,060.
[23] Because this process was conducted prior to the village level randomization and in all villages, regardless of their eventual treatment status, there is no possibility that this method of sample selection could introduce any bias into the evaluation design.
[24] For the first few weeks of data collection only one potential replacement household with a pregnant woman and two potential replacement households with a child under twelve months of age were identified.
[25] As a result of the sampling procedure, different households had different probabilities of being included in the final sample. These sampling probabilities can be used to calculate sampling probability weights, by calculating the inverse of the sampling probability for each household. While our preferred specification will not use these sampling weights, we will ensure that treatment effects are robust to their inclusion.
[26] As was the case for the village sampling, household sampling was done in Stata 14 by Giordano Palloni using the runiform() function to draw a uniform random number for each household.

children's life stage, we elected to carefully balance the number of sampled pregnant women across the two treatment groups.

Because of the time sensitive nature of the treatment, the randomization was done in five separate waves as new data became available. Therefore, it was not possible to observe the number of sampled households with a pregnant woman in all sample villages prior to conducting the randomization. We therefore follow Guiteras et al. (2016) and select a treatment allocation with a high relative efficiency in each wave, taking as given the treatment allocation for villages that were included in any previous waves. For any treatment allocation, the relative efficiency provides a scalar measure of the balance in observable characteristics between potential treatment groups, relative to the most balanced treatment allocation possible given the realized values for all characteristics being considered. A relative efficiency of one—the maximum possible value—implies that the treatment allocation is as balanced as possible for the selected characteristics and observed data.[27] The improved balance in observable characteristics generated by a higher relative efficiency can increase statistical power to detect treatment effects (Guiteras et al., 2016).

For each wave of data, 1,000,000 distinct treatment allocations were generated after restricting the number of treatment villages and control villages in the wave to be as close as possible to equal. Using the observed number of households with a pregnant woman to be sampled in each village, the relative efficiency of the allocation is calculated for each of the 1,000,000 potential allocations. The 1,000 allocations with the highest relative efficiency are retained, and one of these 1,000 most efficient allocations is randomly selected by assigning 1,000 uniform random numbers—one to each potential allocation—and choosing the allocation with the lowest random number draw. The non-selected 999 most efficient allocations are preserved to use for randomization inference during the analysis stage of the evaluation.

In all waves after the first wave of CLE data, the procedure was modified to take as given all previous assignment of villages to treatment groups. That is, for each of the 1,000,000 potential treatment allocations generated for a new wave of villages, the relative efficiency is calculated using the potential treatment allocation for previously unassigned villages but the actual treatment assignment for villages that were allocated to a treatment group during a previous wave of the randomization.

Ultimately, there were five waves of village level randomization with between 3 and 47 villages allocated to treatment groups in each wave. Across all five waves 180 villages were allocated to a treatment group: 90 to the control group and 90 to the treatment group.

3.4.4 Household Level Randomization

While the village level randomization described in Section 3.4.3 is the basis for inference for research questions 1 through 4, it will not generate variation that can be used to answer research question 5:

5. Does sending the mNutrition programme to the mobile phones of both women—either the pregnant woman or the mother of the child under twelve months of age at baseline—and their spouses have a differential impact on the other primary and secondary outcomes?

To answer research question 5 the quantitative evaluation also includes a second, household level randomization among the households in treatment villages that met an additional eligibility criteria. To be eligible for the second randomization, households had to reside in a village assigned to the treatment group, they must have been selected through the household sampling process described in Section 3.4.2, and the primary female and the primary male had to each own a distinct mobile phone.

For the subset of treatment households that met these additional eligibility criteria, we randomly allocated half to receive the mNutrition content on just the mobile phone of the primary female and half to receive the

[27] We refer interested readers to Guiteras et al. (2016) for a more detailed discussion.

mNutrition content on both the mobile phone of the primary female and the mobile phone of the primary male.

Of the 1,428 households successfully interviewed in treatment villages, 276—or 19.3%—were identified as being eligible for the household level randomization. This number is low largely because, as we discuss in more detail below, the enumeration team was able to interview a male respondent in less than half of the baseline sample. Of the treatment households where a male respondent was successfully interviewed roughly 38.7% were deemed eligible for the household level randomization—in line with our pre-data collection expectation for the percent of sample households where both the primary male and primary female would have distinct mobile phones.

While the low eligibility rate will not affect our ability to estimate internally valid treatment effects, it certainly may call into question the external validity of our conclusions. For example, if— as seems likely— households where both the primary male and primary female own distinct mobile phones are different from those where one member does not own their own mobile, we may not be able to generalize the conclusions based on this household-level randomization to the broader population being studied. Still, the results will be informative for nearly 20% of our study population, a meaningfully large and independently interesting sub-group. Further, this has no bearing on our ability to answer the primary research questions of the evaluation, which the study was designed to answer.

138 of these eligible treatment households were assigned to receive the content on the mobile phones of both the primary female and the primary male (T-F+M) and the remaining 138 were assigned to receive the content on just the mobile phone of the primary female (T-F).

The household level randomization will generate experimental variation in the information set of the primary male respondent in eligible households. This creates an opportunity for us to learn about how information flows within households and how household members form beliefs, both of which are critical open questions for policy makers that hope to generate behaviour change through the provision of information.

If spouses share all information with one another and view the information they receive from their spouse regarding health and nutrition as credible, we should expect to see no differences in outcomes between treatment households assigned to the T-F+M group and those assigned to the T-F group. The primary female, upon receiving an mNutrition message will impart the content to her spouse, they each will decide whether to change their beliefs in response to the information they received, and in accordance with some household decision-making process the pair will determine whether to change their behaviour.

However, it need not be true that there are no within-household frictions that affect information sharing or receipt between spouses. For example, spouses may, intentionally or unintentionally, not convey the content of a message or group of messages with one another. As a result, the spouse that did not directly receive the message on their mobile may be less likely to adjust their beliefs, and the outcome of the household decision-making process may be different from what it would have been if both spouses had received the information directly.

A second reason why outcomes may differ between the T-F+M and T-F groups is that spouses may not view one another as credible sources for health and nutrition information. In particular, it is possible that the primary males in the sample of treatment households eligible for the household level randomization are less likely to change their beliefs in response to information given to them by their spouse than they are in response to information sent to them via SMS by an NGO-type organization. For households where this is true, even if intrahousehold information sharing is complete, beliefs, behaviours, and outcomes may be changed more by the T-F+M treatment than by the T-F treatment. The household level randomization will enable us to test for the presence of some type of information sharing friction within the household, either incomplete communication or credibility issues between spouses.

3.4.5 Household Survey

The household survey is the principal source of information for the primary and secondary outcomes. In addition, the questionnaire collects data on indicators expected to be strongly correlated with the primary and secondary outcomes—which can be used to help improve the statistical precision of the treatment effect estimates—as well as on measures that are important for testing the different causal mechanisms that could generate differences in the final outcomes of interest. When paired with the endline data collection in the context of the randomized evaluation design, the baseline instrument will enable us to carefully test for causal effects of the mNutrition programme at different levels of the causal chain. Data collected during the baseline survey—GPS coordinates and all mobile phone numbers for each household—will also be used to help locate the households surveyed at baseline for endline fieldwork.

The baseline questionnaire elicited information from all households included in the sample, regardless of their treatment group. When possible, the household survey collected information from three members per household: the primary female, the primary male, and the child under twelve months of age.[28] While every household in the sample had a primary female—either the pregnant woman or the mother of the child under twelve months—not all completed interviews contain information from a primary male respondent, and only 67% of surveyed households had a child under twelve months of age.

The full text of the baseline questionnaire can be found in Annex B. The survey elicited information on the following categories: geographic coordinates and basic household identifiers; household composition and demographic characteristics of household members; household assets and physical characteristics of the household structure; general health and vaccination history for children under the age of five; HIV/AIDS awareness of the primary female respondent; marriage and fertility history of the primary female respondent; desired future fertility for the primary female and the primary male respondents; mobile phone access and usage of the primary female and primary female respondents; antenatal and prenatal care for children under the age of three; knowledge and beliefs about infant and young child feeding practices (IYCF) for the primary female and primary male; IYCF practices for the two youngest children under the age of five; women's dietary diversity for the primary female; primary female trust in nutrition and health information from different sources; and direct measures of height and weight for the child under twelve months of age.

3.4.6 Treatment Offer

While the mNutrition content is available, free of charge, to households across Tanzania, we expect that take-up of the programme will be low in Iringa region in the absence of the evaluation activities. Largely, this is because the promotion of the previous programme operated by the mHealth Tanzania Public-Private Partnership (mHealth Tanzania-PPP) was dependent on relationships with health clinics and partner organizations operating within certain regions to help register potential users, and they had no existing relationships with organizations operating in Iringa; this is, in part, why Iringa was selected as the study location. In addition, due to resource constraints it was clear that in-depth marketing of the mNutrition content would not be feasible for the implementing partners to undertake in all regions of Tanzania.

Though the above described measures should help to ensure that take-up of the mNutrition programme is low in control villages, to precisely estimate the treatment effects of mNutrition on outcomes we also need take-up of the programme to be high among sampled households in treatment villages. To increase participation among treatment households, we relied on door-to-door offers of access to the programme by the OPMT field team. After completing the baseline household survey in treatment villages, enumerators were instructed to read the following text to the primary female respondent: "Thank you for your participation in the study. We are now done with the interview. Your household has been selected to receive free text messages that contain health and nutrition information that may benefit you and your

[28] Anthropometry measurements were recorded for all children under the age of twelve months, conditional on receiving consent from the child's mother or father. No additional information was collected from this child.

children. This service is called mNutrition and it is provided by Wazazi Nipendeni SMS service. Are you willing to give consent for your household to receive these text messages?" If the primary female respondent gave her consent, then the mobile number which she indicated was her "Main phone number" and her gestational age (if pregnant) or the age of her child under twelve months were recorded. Similarly, in treatment households where the primary male respondent was also randomly selected to receive the mNutrition content, he was also asked for his consent to receive the programme on his main mobile number. The enumeration teams were instructed not to mention the mNutrition programme until after the full household survey had been completed, to reduce the likelihood that knowledge of the programme or an expectation of future benefits would affect the answers given by respondents.

OPMT and IFPRI processed the baseline survey data on a rolling basis to extract the information necessary to register households for the mNutrition programme as early as possible. The interview date, age or gestation at the time of the interview date, and the date the data became available were used to update the expected gestation or age of the child so that they were accurate when the messages began being sent to beneficiary households.

3.5 Empirical Strategy

3.5.1 Intent to Treat (ITT) Estimates

Because the offer of treatment is randomly assigned, we will use the systematic variation in take-up of the mNutrition programme generated by the random offers to identify the causal impact of the programme. The random assignment of the treatment offers ensures that unbiased estimates of the treatment effects can be estimated using simple differences, difference-in-differences, or Analysis of Covariance (ANCOVA) specifications because observable and unobservable characteristics of children, households, and communities will be balanced across treatment and control villages. However, ANCOVA models, which control flexibly for a baseline measure of the outcome, are likely to be the most efficient of the three estimators, particularly when the autocorrelation for the outcome being considered is low (McKenzie, 2012). Therefore, for panel outcomes—those that are observed at both baseline and endline—we will use ANCOVA to generate our primary estimates, and simple differences and difference-in-differences models as robustness checks.

For outcomes that were not observable at baseline, children's dietary diversity for example, we will use a simple differences specification to estimate treatment effects. To identify the simple differences treatment effects we will estimate the following Ordinary Least Squares (OLS) regressions:

$$Y_{1ihv} = \beta_0 + \beta_{1,SD} Treatment_v + \delta_x X_{0ihv} \quad + \varepsilon_{ihv} \,,$$

where Y_{1ihv} is the outcome measured at endline ($t = 1$), for individual i, in household h, in village v, β_0 is a constant term, $Treatment_v$ is an indicator equal to one if village v was randomly assigned to the treatment group, X_{0ihv} is a vector of observable characteristics for individual i measured at baseline ($t = 0$), and ε_{ihv} is an error term which we will cluster at the village level. In this model, δ_x represents the vector of coefficients on the controls and $\beta_{1,SD}$ is the parameter of interest: the simple differences causal effect of being offered access to the mNutrition programme on the outcome.

Though the random assignment of treatment to villages ensures that $\beta_{1,SD}$ is an unbiased estimate of the causal effect of the mNutrition offer, if a baseline measure of the outcome is available there are more efficient methods available. In particular, the difference-in-differences specification can be estimated by running the following OLS regression:

$$Y_{tihv} = \sum_{t=0}^{1} \delta_t + \beta_{1,DD} Treatment_{tv} + \beta_2 EverTreat_v + \delta_x X_{tihv} \quad + \varepsilon_{tihv}$$

where $t \in \{0,1\}$, Y_{tihv} is the outcome measured in period t the δ_t are time period fixed effects, $EverTreat_v$ is an indicator for whether village v ever receives the mNutrition programme offers, $Treatment_{tv}$ is an indicator for whether village v received offers for the mNutrition programme in period t, X_{tihv} is a vector of time varying controls, and ε_{tihv} is an error term which, again, will be clustered at the village level. For outcomes with an autocorrelation above 0.5, the difference-in-differences estimator will yield more precise estimates of the treatment effects ($\beta_{1,DD}$) than the simple differences model.

However, with one baseline survey and one endline survey, treatment effects estimated through an ANCOVA specification will have lower variance[29] than the difference-in-differences treatment effects as long as the autocorrelation for the outcome is below one.[30] ANCOVA treatment effects can be estimated by running the following OLS regression:

$$Y_{1ihv} = \beta_0 + \beta_{1,A} Treatment_v + \beta_Y Y_{0ihv} + \delta_x X_{0ihv} \quad + \varepsilon_{ihv},$$

where all of the parameters are defined in the same way as in the above simple differences equation and Y_{0ihv} is the baseline measure of the outcome. In addition to providing more efficient estimates of the treatment effects, the ANCOVA model will allow us to estimate the relationship between the baseline and endline measures of the outcome. For outcomes that are observable at both baseline and endline, we will rely on this ANCOVA specification to generate the main treatment effect estimates.

3.5.2 Local Average Treatment Effects (LATE) for Compliers

As mentioned in Section 1, the specifications described above enable us to estimate intent to treat (ITT) treatment effects; that is, the point estimates capture the impact of the random offer of access to the mNutrition programme on outcomes. However, under two assumptions[31] we can also estimate the local average treatment effect (LATE) of receiving the mNutrition content for compliers: households that were induced to register for the programme by the randomly assigned door-to-door offer.

The first required assumption is that the mNutrition offer only affected outcomes indirectly, by increasing the likelihood that households received the mNutrition content on a mobile phone. Given the brief nature of the treatment offer—enumerators were instructed to offer households access to "free text messages that contain health and nutrition information that may benefit you and your children"—it is quite likely that the assumption will hold in this context; the prospect that this momentary interaction could directly affect knowledge, beliefs, behaviours, or nutrition outcomes is unlikely.

The second assumption necessary for estimating LATE for compliers is that the randomly assigned offer of access to the mNutrition programme does not decrease the likelihood that any household or household member registers to receive the content. This assumption would be violated if, for example, the enumerator introduction to the programme were so ineffective that it convinced households who otherwise would have discovered and registered for the content, not to give their consent to receive the mNutrition messages. In practice, it is hard to envision a scenario where the offer of access to the mNutrition content could make a household less likely to participate in the programme. Therefore, the second assumption is also likely to be satisfied in the context of this evaluation.

If both assumptions are satisfied, the LATE for complier households can be estimated through Two Stage Least Squares (2SLS) estimates of the ANCOVA or simple difference models discussed in the previous subsection, using the random assignment of villages to receive the mNutrition treatment offer as an

[29] Though we know the variance of the treatment effects estimated through ANCOVA will be lower than the variance for those estimated with simple differences or difference-in-differences, we cannot know the size of the difference between the variances until we have collected endline data and estimated the autocorrelations for each outcome.

[30] See McKenzie (2012). In an experiment with one baseline and one follow-up survey, the ratio of the variance of the difference-in-differences estimate to the variance of the ANCOVA variance is given by $\frac{2}{(1+\rho)}$, where ρ is the autocorrelation.

[31] See Imbens and Rubin (2015) for a complete discussion.

excluded instrument for observed take-up of the programme.[32] Specifically, we will estimate the following models:

$$Y_{1ihv} = \beta_0 + \beta_{1,2SLS}\widehat{TT_{hv}} + \beta_Y Y_{0ihv} + \delta_x X_{0ihv} \quad + \varepsilon_{ihv};$$

$$TT_{hv} = \gamma_0 + \gamma_{1,A} Treatment_v + \gamma_Y Y_{0ihv} + \gamma_x X_{0ihv} \quad + u_{ihv};$$

where $Treatment_v$ is the indicator for whether the household resides in a village v that was assigned to receive the mNutrition offer, and TT_{hv} is an indicator for whether household h in village v actually registers for the mNutrition programme. $\widehat{TT_{hv}}$ is the predicted value for the take-up of household h in village v from the take-up equation (the second equation listed above). In this context, $\beta_{1,2SLS}$ represent the estimated effect of receiving the mNutrition content for the sub-sample of households that are induced to participate in the programme by the randomly assigned offer.

The LATE treatment effects for compliers provide a potentially more policy relevant parameter than the previously discussed ITT effects; they represent, albeit for a specific sub-population (compliers), the causal effect of exposure to the mNutrition messaging. Regardless, we will estimate and discuss both parameters in order to provide more complete conclusions about the causal effects of the mNutrition programme.

As with any longitudinal study, the attrition of households between baseline and endline poses a potential problem. We will mitigate this risk by collecting detailed GPS information and a list of every mobile phone number owned for each household in our sample. When endline fieldwork begins, we will use this information to help us ensure that: 1) we have the correct residential location for all households in the sample, and 2) the households are available on the planned day for their endline interview. Undoubtedly, there will be still be some attrition between the two survey rounds, though we do not expect any differential attrition by treatment arm. As a result, we will use baseline data to predict endline participation, generate endline survey completion weights (the inverse of the predicted probability of completing the endline survey conditional on baseline characteristics), and ensure that our results are robust to including these weights in the above listed empirical specifications.

3.5.3 Statistical Inference

With 180 study clusters included in the quantitative evaluation, statistical inference based on cluster-robust standard errors is likely to be valid and result in tests of the correct size (Bertrand et al., 2004). However, for each outcome we will also conduct randomization inference as a robustness check on the tests that rely on the asymptotic normality of a test statistic based on a finite sample (Fisher, 1935; Rosenbaum, 2002; Greevy et al., 2004; Imbens and Rosenbaum, 2005; Small et al., 2008). Randomization inference offers a non-parametric alternative for testing the sharp null hypothesis of no treatment effect for any household.[33]

To conduct randomization inference for the sharp null hypothesis of no treatment effect for any household, we will proceed as follows. First, we calculate the test statistic, for example the difference in mean outcomes between villages actually assigned to the mNutrition treatment group and villages assigned to the mNutrition control group, $\hat{\beta} = \bar{y}_{ihv,T=1} - \bar{y}_{ihv,T=0}$. Next, we conduct $R = 100{,}000$ placebo treatment assignments. That is, 100,000 different times we re-assign 90 of the 180 sample villages to a placebo treatment group and the other 90 to the placebo control group. For each repetition r, we calculate the treatment effect, which under the null of no effect for any household is simply $\hat{\beta}_r = \bar{y}_{ihv,T_r=1} - \bar{y}_{ihv,T_r=0}$. Here $T_r = 1$ denotes that the household resides in a village that was assigned to the placebo treatment group in repetition r, $T_r = 0$ indicates that the village was assigned to the placebo control group in

[32] For the linear models specified in this context, the 2SLS estimates of LATE for compliers will be equal to the ratio of the ITT estimate of the treatment offer on the outcome to the ITT estimate of the treatment offer on take-up of the programme.
[33] In practice, Randomization Inference can be used to test the sharp null of a treatment effect of any size, not just no treatment effect.

repetition r, and by the null of no effect for any household $y_{ihv,T_r=1} = y_{ihv,T_r=0}$, so the observed outcome does not need to be adjusted.

After performing the placebo treatment assignment 100,000 times, we will have a distribution of test statistics from all of the repetitions: $\{\hat{\beta}_r\} = \{\hat{\beta}_1, \ldots, \hat{\beta}_R\}$. To assess the plausibility of the observed test statistic ($\hat{\beta}$) under the null hypothesis, we calculate the share of repetitions for which $\hat{\beta}_r > \hat{\beta}$. This share is an empirical p-value for the sharp null of no treatment effect for any household.[34]

[34] Though the exact p-value could be computed by generating all $\binom{180}{90}$ possible treatment assignments, this is computationally infeasible. With 100,000 repetitions, the empirical p-value will have a standard error less than or equal to $\frac{1}{2\sqrt{R}} = 0.0016$.

4 Baseline Data Collection

The household survey questionnaire was composed to collect information on potential outcomes, basic demographics, indicators that were likely to be predictive of the potential outcomes, and intermediate outcomes that are relevant for testing different causal mechanisms. The full text of the household questionnaire can be found in Annex B. To measure outcomes, the indicators relevant for testing potential causal mechanisms, and important demographic characteristics, the questionnaire contains a wide array of survey questions, enumerator measurements, and enumerator observations. The IFPRI team exerted considerable effort to ensure that the full household survey could be completed in less than an hour and thirty minutes, to reduce the likelihood that respondent fatigue would increase measurement error to the survey indicators.

OPMT translated both questionnaires from English into Swahili. The translation was undertaken by the enumeration teams, under the supervision of the fieldwork managers. All members of the enumeration team were fluent in both Swahili and English, and all had past experience conducting and translating questionnaires into Swahili. In addition, as a part of the piloting process (described in more detail below), the Swahili versions of the translated questionnaires were tested and further refined to make sure that the final translations would be interpreted correctly by respondents.

4.1 Ethics Approval

IFPRI received approval from their Institutional Review Board (IRB) for the Tanzania quantitative evaluation design described in Section 3. The letter of authorisation is included in this document as Annex C**.** An application for the qualitative and quantitative surveys was submitted to and approved by the Commission for Science and Technology (COSTECH) in Tanzania and a research permit was granted (see Annex D). This application was based on the original encouragement design so the research team have checked this with OPMT and submitted details on the modifications to the design as part of the COSTECH permit renewal process in July of 2017. The original encouragement design was adjusted in response to the mNutrition content being offered for free. The research team determined the best way forward was to identify a study region where mHealth Tanzania would not promote the WN product and where baseline take-up of the existing “Healthy Pregnancy, Healthy Baby” programme was low.

As an overall guiding principle, the research team sought to conduct itself in a professional and ethical manner throughout the initial exploratory study, with respect for integrity, honesty, confidentiality, voluntary participation, impartiality and the avoidance of personal risk. These principles were guided by the OECD (2010) DAC Quality Standards for Development Evaluation and DFID’s (2011) Ethics Principles for Research and Evaluation, which will be followed for the duration of the evaluation.

National-level ethical approval for the quantitative baseline was obtained from the Tanzania Commission for Science and Technology (COSTECH) in July 2016 (prior to the start of the fieldwork). Ethical review was also sought from the IDS Research Ethics Committee and was obtained in September 2016, also prior to the start of the fieldwork.

Informed, written or oral, consent was collected from all participants prior to the start of the interviews. The entire field team was trained on ethical data collection and signed an ethical conduct form prior to the start of the field work. Participants received two bars of soap as compensation for their participation in the interviews.

Confidentiality of the data is protected by recording survey interview responses using Computer Assisted Personal Interviewing (CAPI), so no hard copy versions of survey questionnaires are available. All files containing raw and analysed data are securely stored in password-protected databases. Access to the complete data is restricted to the IFPRI/IDS/Gamos evaluation team. A unique household ID is assigned to each household. The name and geographic location of the respondent will be kept in a separate data file to

which only the research team will have access. Anonymized versions of the data sets that exclude these personal identifiers will be the ones made available for public access.

4.2 Enumeration Team and Trainings

Prior to beginning the formal enumerator training, the OPMT team leaders, fieldwork managers, and members of the IFPRI quantitative team pre-tested the survey instrument in the Bagamoyo District of Pwani Region Tanzania during the week of September 12th, 2016. The purpose of the pre-testing was to identify questions that were poorly understood by rural households, highlight Swahili translations that were not interpreted in the desired way by respondents, and to make sure that the duration of the household survey was less than an hour and a half.

Following the completion of pre-testing, OPMT staff programmed the updated questionnaire in Survey Solutions,[35] a computer-assisted personal interviewing (CAPI) software created by the Development Research Group at the World Bank. OPMT and IFPRI staff tested the CAPI version of the questionnaire at length to check that answer codes and skip patterns were correct for each question, that all questions appeared if and when they should, and to make sure that the software performed without significant issues (e.g. programme crashes were rare). The CAPI programme was then uploaded onto tablets for the beginning of the enumerator training.

Enumerator trainings for the baseline data collection activities were conducted in Dar es Salaam, Tanzania and in Iringa Town, Iringa Region, Tanzania between the 3rd and 14th of October, 2016. The trainings were led by five senior staff members from OPMT and a senior research analyst from IFPRI. The data manager from OPMT was also present throughout the trainings. The first week of training took place in Dar es Salaam, before moving to Iringa Town for the second week. Anthropometry specialists were also trained during this two week period.

The CLE enumeration team consisted of two teams with five enumerators each, one of whom was also identified as the team supervisor. The CLE teams completed village listings in two different villages per field day, and they completed the listing activities in all 180 villages in 45 days.

The household survey enumeration team included four teams with six enumerators, one supervisor, and one anthropometric specialist per team. Each enumerator completed, on average, three household interviews per day, and the full household survey data collection was completed in 45 days.

The OPMT data collection team was careful to ensure the quality of the data collection. This was done primarily in four ways. First, team supervisors travelled with the enumeration teams, sat in on interviews, and reviewed the data being collected. Second, a data manager and fieldwork manager also were present for the first week of household survey fieldwork, during which time also sat in on household interviews, checked the data being recorded, and offered additional feedback to enumerators. Third, the fieldwork manager and data manager conducted village revisits where they elicited feedback from village leaders and respondents; they also made follow-up phone calls to verify data in some cases. Fourth, the OPMT programmed the CAPI such that any data inconsistencies would be highlighted immediately for the enumerator.

The baseline data analysis was conducted by Natasha Ledlie of IFPRI in Washington, D.C., using Stata 14 between January and July of 2017.

[35] More information on Survey Solutions can be found on their website: https://solutions.worldbank.org.

4.3 Fieldwork Experiences

4.3.1 Achieved Sample Size and Interview Completion Rate

4.3.1.1 Community Listing Exercise

The CLE data collection began the 3rd of October, 2016 and was completed November 18th, 2016. Listing interviews were conducted in Swahili and enumerators attempted to identify the most responsible household member available at the time of the visit to respond to the CLE questions.

Prior to beginning a CLE questionnaire, enumerators were instructed to ask for the consent of the respondent, who typically was a household member over the age of fifteen. The enumerator only completed a CLE for the household if they received verbal consent to participate in the CLE study from the potential respondent.

The CLE data collection team successfully interviewed 23,592 households in 180 villages between October 3, 2016, and November 18, 2016. Of those 23,592 households, 1,568 households with a pregnant woman and another 2,885 households with a child under twelve months of age (and that child's mother) were identified as being eligible to participate in the study.

Table 4.1: Basic Statistics from the CLE

	Total interviews
Total households interviewed	23,592
Identified with a pregnant woman	1,658
Identified with a child under 12 months	3,044
Of those with a pregnant woman or a child under 12 months who had a mobile phone	4,466
Of those with a pregnant woman or a child under 12 months who were literate in Swahili and had a mobile phone	4,453

4.3.1.2 Baseline Household Survey

Data collection for the baseline household survey was conducted by OPMT in the Iringa region of Tanzania between October 17 and December 10, 2016. In that time, interviews were completed with 2,833 households—1,405 households in control villages and 1,428 households in treatment villages.

Table 4.2: Basic Statistics from Household Survey

	Targeted interviews	Actual interviews	Percent
Identified with a pregnant woman	1,080	948	87.8%
Identified with a child under 12 months	1,980	1,885	95.2%
Total	3,060	2,833	92.6%

Before beginning a household survey, enumerators read the respondent a brief description of the study that was being conducted, informed them that their participation in the study was voluntary and that they could discontinue participating at any time, and asked whether they agreed to respond to the household interview questions. The full written consent can be seen at the beginning of the household survey in Annex B. This process was repeated with both the primary female and primary male respondents. In all, only two females and sixteen males did not consent to participate in the study.

Up to five replacement households were randomly selected in each village (two households with a pregnant woman and three households with a child under 12-months). These replacements households were ranked and enumerators were instructed to use replacement households in the order they were listed on the household sample list from that village. 260 replacement households completed household interviews (9.2%). Of these replacements, 76 were households with a pregnant woman and 184 were households with a child under 12-months of age.

Including replacements, the enumeration team completed 92.6% of the total targeted number of baseline interviews, 87.8% of the targeted number of baseline interviews with a pregnant woman, and 95.2% of the targeted number of baseline interviews with the mother of a child under twelve months of age. There were a number of reasons why these completion rates were below 100%, but the most limiting factor was the lower than expected fertility rates observed in rural Iringa that we discuss in Section 3.4.2. In addition, the enumeration team attempted to replace 251 originally sampled households when there was no longer a working mobile phone (i.e. the mobile had been lost or broken between the CLE and baseline visit), 63 households where children had aged out of sample eligibility (the identified child was now older than one year of age), 11 where the household was not available to interview on the day the data collection team came to conduct the baseline survey, 43 where the targeted woman or child was determined to not be a permanent household member (this often occurred when women who travelled to their parents' home to give birth were identified as household members during the listing).

4.3.2 Absenteeism of Primary Male Respondents

While the baseline household survey successfully completed interviews with 92.6% of the targeted number of households, interview completion rates were low for the targeted primary male respondents. Across all 180 study villages, enumerators successfully interviewed a male respondent in just 49.5% of the sample households.

The male absenteeism rates were likely driven by two primary factors. First, it was not necessary for the CLE to identify a primary male respondent in the household for the household to be deemed eligible to participate in the study. As a result, 1,415 households in the baseline sample did not have a primary male respondent—for example, there were multiple instances of single mothers, widows, and households where the primary male was a long-term migrant and not currently residing in the household. As measurement of the primary outcomes was not contingent on being able to interview a primary male, IFPRI elected not to require that there be a primary male in the household for that household to be eligible to participate in the study.

Second, the baseline data collection only allowed for one field day in each of the 180 sample villages. As a result, there were 1,415 instances where the individual who would have been the primary male respondent was out of the household, typically working on a distant plot of land, or there was no male household member eligible to be the primary male respondent. In cases where a respondent was working but not far from the home, enumerators attempted to interview the person in the alternate location; however, because many households in the sample worked land that was located hours away from their home, in many instances it was not practical for the enumerators to travel to the plot in order to complete the interview with the primary male.

While unfortunate, this low interview completion rate among males should not impact the quantitative evaluation plans. None of the primary outcomes require that information be collected from the primary male household member. That said, a higher number of completed primary male interviews would have improved our ability to explore how information flows within households, by increasing the statistical power for detecting differential treatment effects between households receiving the mNutrition content on just the primary female respondent's mobile phone as compared to households receiving the mNutrition content on the mobile phones of both the primary male and the primary female. As we discuss above, by changing the sample of households eligible for the household-level randomization, the results based on this randomization may also be less representative of the effect of providing multiple household members with

the mNutrition information on outcomes for rural households in Iringa. However, they will still capture the causal effect for a non-trivial share of the population in rural Iringa.

4.3.3 Consent for mNutrition in Treatment Households

The evaluation design relies on a random offer of access to the mNutrition programme to drive differences in registration across treatment and control villages. In all treatment households, the primary female respondent was given a brief description of the mNutrition content[36] and asked whether she was willing to give her consent for the household to receive the messages on a mobile phone. She was also informed that the messages would be free of charge to the household. The mNutrition description and request for consent was always conducted at the end of the household survey. No questions were asked after introducing mNutrition to the respondent. The process for the primary male respondent—if he was able to be interviewed—was exactly the same as for the primary female: a brief description was read to the respondent, at which point he was asked for his consent to receive the mNutrition messages.

Few interviewed respondents were not willing to consent to receiving the mNutrition content. Of the 1,428 primary female respondents in treatment villages asked for consent, 1,421—or 99.5%—gave their consent to receive the content on their mobile. Similarly, of the 713 primary male respondents in treatment villages asked for consent, 712—or 99.86%—gave their consent to receive the content on their mobile. In total, 2,133 of the 2,141 interviewed treatment individuals—or 99.62%—were willing to agree to be registered for the programme.[37] At least one number from each of these households was sent to Wazazi Nipendeni to be registered for the programme.

While we cannot possibly know what the actual difference in take-up between treatment and control households will be until the endline, the high take-up among treatment households suggests that the quantitative evaluation may be better powered to detect statistical differences in the outcomes than originally anticipated, despite not reaching the targeted number of completed interviews. For example, even if 10% of control households register for the programme during the study period and attrition among those households successfully interviewed at baseline is 10%, the evaluation should be powered to detect differences in HAZ scores of around 0.2 standard deviations between treatment and control households; smaller than the 0.25 standard deviation differences the evaluation was designed to detect. Further, if take-up rates in control villages stay lower than 10% or attrition is lower than 10%, the evaluation will be able to detect even smaller differences in HAZ.

4.3.4 Health Centre Referrals for Severe Acute Malnutrition

Height and weight measurements were taken for children under twelve months of age at baseline. For children that were identified as being severely acutely malnourished—a weight-for-height z-score below -3—the CAPI programme prompted the enumerator to refer the child to a health facility. Enumerators briefly described the situation to the mother or father of the child and advised the parent to take the child to a health facility. The parent was also provided with a referral form to bring to the health facility.[38] Of the 1,851 children with non-missing height, weight, age, and sex, just eight were identified as severely acutely malnourished.

[36] "Thank you for your participation in the study. We are now done with the interview. Your household has been selected to receive free text messages that contain health and nutrition information that may benefit you and your children. This service is called mNutrition. Are you willing to give consent for your household to receive these text messages?"

[37] Primary females in seven control households were also asked for their consent to receive the mNutrition messages along with primary males in five of those seven households. While all individuals from those seven households gave their consent, the IFPRI team only registered phone numbers that belonged to households in treatment villages for the programme. Where therefore expect that these seven households will not receive the content during the study period.

[38] As a part of the fieldwork, team supervisors also visited local health facilities to alert them that they would be referring severely acutely malnourished children in the area.

5 Baseline Data: Sample Characteristics and Balance

Baseline data were successfully collected from 2,833 households. In this section, we discuss the data with two primary goals: characterizing the observable attributes of the household sample, especially those relevant to the mobile based information intervention being evaluated and assessing balance in baseline characteristics across the two village level treatment groups. We do this by presenting the means and standard deviations from baseline data for the full sample, as well as disaggregated by village treatment assignment. Each table of baseline data is followed by a brief discussion of the statistics just presented, and their implications for the quantitative evaluation.

Balance in baseline characteristics across the two treatment groups is central to the success of the RCT evaluation strategy. Imbalance in observable attributes at baseline, especially those thought to be strongly correlated with the outcomes of interest, typically casts doubt on the ability of the evaluation to identify the causal effect of the intervention being investigated. However, there is no clear consensus in the evaluation literature about how best to determine whether there is sufficient balance, or overlap, in the distribution of a characteristic across the treatment and control group. In particular, while a basic comparison of the mean of a variable in the treatment group with the mean of that same variable in the control group can be informative, it is hard to draw any sound inference about whether the observed difference is meaningful. We elect to present two sets of balance measures for each baseline characteristic: the p-value from a t-test of a null hypothesis that there is no difference in means between the two treatment groups and the normalized difference suggested by Imbens (2015).

5.1.1 P-Value from a Test of No Difference in Means

Most quantitative evaluations resort to statistical tests for the equality of means across treatment groups; in effect, these test statistics capture how large the differences in means are relative to the typical variation in a variable observed in the data. We follow this practice by presenting, for each baseline characteristic, the probability (p-value) of observing a difference in means between the treatment and control group that is at least as large as the observed difference, given that the null hypothesis of no difference between the two groups is true. This p-value is computed based on a t-test of the null hypothesis that there is no difference between the two groups, from a regression of the characteristic on an indicator for whether each household resided in a treatment village at baseline with clustering standard errors at the village level. A successful randomization should lead to few statistically significant differences in observable characteristics between the two groups.

Because we test for differences across the two treatment groups for many different baseline characteristics, even if the randomization was successful, we will observe some statistically significant differences. For example, interpreting characteristics based on the convention that a p-value below 0.05 is significant, we should expect to observe a significant difference for one out of every twenty tests simply by chance.[39] However, observing a significant difference for substantially more than one out of twenty tests would indicate that the randomization was not successful, and suggest that any differences in outcomes at endline could be attributable to the baseline imbalance, rather than the mNutrition treatment.

5.1.2 Normalized Difference

Though assessing balance in observable characteristics by calculating p-values from a test of the null hypothesis of no difference between the treatment and control group is undeniably useful, it is also

[39] The number of significant differences we should expect to observe by chance is actually greater than one out of twenty. This is because, when testing multiple hypotheses simultaneously, the probability of observing at least one difference that is significant at the 5% level is actually greater than 5%. While methods have been developed to adjust for multiple hypothesis testing, we elect to present the unadjusted p-values and instead encourage readers to exert caution to avoid overinterpreting any significant differences.

sensitive to the sample size. Because the p-value is based on the t-statistic—the ratio of the difference in means between the two groups to the standard error for that difference—p-values decrease quickly with the sample size. Therefore, particularly for large sample sizes, large t-statistics and the corresponding low p-values may be less informative about observable balance. We therefore also follow Imbens (2015) and present the normalized difference for each characteristic. The normalized difference is the difference in means between the two groups scaled by the average of the within group standard deviations. Specifically, for characteristic x, the normalized difference is given by:

$$\Delta_x = \frac{\mu_T - \mu_C}{\sqrt{(\sigma_T^2 + \sigma_C^2)/2}},$$

where μ_T and μ_C are the sample means for households in the treatment and control group and σ_T^2 and σ_C^2 are the conditional within-group sample variances for characteristic x, respectively. Like the p-value from a t-test of no difference between the two treatment groups, the normalized difference is scale free (i.e. the difference is calculated relative to the "normal" variation in the variable as measured by the variance). However, the normalized difference is also substantially less sensitive to the sample size: the t-statistic is approximately equal to the normalized difference multiplied by the square root of the total sample size. We therefore use the normalized differences as our preferred measures of balance, and follow Imbens (2015) in interpreting normalized differences below 0.25 as being indicative of baseline balance.

The remainder of this section describes the baseline data and balance in observable characteristics between the treatment and control groups. We begin with a table for basic household demographics including the age, sex, marital status and educational attainment of household members, the physical structure of the household and the sources of sanitation and drinking water, household asset ownership, and Progress out of Poverty Scores (PPI).[40] Next we present information on each of the primary outcomes and other measures of child anthropometry followed by IYCF knowledge and beliefs (the secondary outcome), information sources for health and nutrition information, self-reported mobile phone access and usage data, general child health and vaccination histories for all children under the age of five, antenatal and postnatal care utilization for children under the age of three, awareness of HIV/AIDS, marriage and fertility histories and desired future fertility for the primary respondents.

5.2 Household Demographics, Physical Structure, Amenities, and Wealth

Table 5.1 presents sample characteristics for household demographics, the physical structure of the home, amenities, and household wealth. To measure household wealth, we rely largely on asset indices and poverty indices to help summarize household responses to a large set of asset and wealth-related questions. For the asset indices, we generate indicators for whether the household owned at least one of each asset in that class (e.g. sheep and goats would be included in the livestock index). Principal component analysis was then used to identify the first orthogonal component—the linearly independent component that explains the highest fraction of the total variance in the class—and that component is used as the index for that asset category. We produce asset indices for household consumer durable assets,[41] household production assets,[42] household livestock assets,[43] and a total asset index.

The progress out of poverty index (PPI) uses a country and year-specific set of ten questions to calculate the likelihood that a household is living below different national and international poverty lines. In Tanzania,

[40] See Annex G and http://www.progressoutofpoverty.org/ for a more detailed description of the methods.
[41] The assets included in the household consumer durable asset index are stove/gas burner, metal cooking pots, beds, armoire/cabinet, table, chair, electric fan, lantern, iron (charcoal or electric), audio cassette/cd player/radio, video/dvd, sofa, television, refrigerator/freezer, sewing machine, bicycle, moped, motor vehicle, and boat. See Annex H for details.
[42] The assets included in the household production asset index are saw, solar energy panel, plough & yoke for animals, hoe, spade/shovel, reaper/sickle, manual sprayer, rake/harrow, wheelbarrow, hauling carts, tractor, water pumping set, fertilizer distributor, spraying machines (for chemicals/fertilizer), and harvesting and threshing machines. See Annex H for details.
[43] The assets included in the household livestock asset index are bulls/oxen/cows, sheep and goats, chicken and other fowl, and donkeys and mules.

the ten questions used to generate the PPI score based on the most recent PPI (from 2011) are the age distribution of household members, whether all household members aged 6-18 are in school, the materials used to construct the roof and walls of the household structure, the primary fuel used for cooking, and ownership of televisions, radios/cassette or tape recorders, lanterns, and tables, and whether the household cultivates crops and owns livestock.

On average, households have 5.3 members in the overall sample, 5.2 members in control villages, and 5.3 members in treatment villages. Household heads are on average 38.6 years of age, 38.5 in control villages and 38.7 in treatment villages, while the primary females are 27.0 years of age on average in the overall sample, 26.8 years of age control villages, and 27.1 years of age in treatment villages.

Nearly all household heads are married, with 74.2% of the household heads in the overall sample being in monogamous marriages and another 13.9% being in polygamous marriages; the treatment and control averages for household head marital status closely mirror those for the full sample. Similarly, the marital status patterns for the primary female are nearly identical in both treatment groups: roughly one fifth have never been married, around 67.7% are currently in a monogamous marriage, and the remaining 8.9% are in polygamous marriages. Of those women in polygamous marriages, 30.6% are the first wife, another 59.8% are second wives, and the remainder are third or higher order wives. Among married women, their existing marriages started, on average, 88.8 months ago in the overall sample, 86.5 months ago in the control group, and 91.0 months ago in the treatment group.

Crop production is the primary activity for both heads of household and the primary females in the sample. For heads of household, 73.1%, 72.8%, and 73.5% are primarily involved in crop production in the overall sample, in control villages, and in treatment villages. For the primary females, the corresponding figures are 76.8%, 76.2%, and 77.4%. Among school aged children (ages 5-20), 66.2% in the overall sample are currently attending school; 65.6% and 66.8% are attending school control villages and in treatment villages. 93.0% of household heads and 93.6% of the primary female respondents have ever attended at some formal school.

Of the1,928 households with a child under twelve months of age (a focus child), 51.2% of the children under twelve months of age are female (have a female focus child) in the full sample, 53.1% are female in control villages, and 49.4% are female in treatment villages.

The PPI scores are extremely balanced across the two treatment arms: 39.5 in control villages, and 39.2 in treatment villages. A PPI score of 39 corresponds to a household having a 58.1% chance of being below 150% of the national poverty line in Tanzania, and a 77.2% chance of living on less than $2.00 per day in 2005 US dollars. For poverty lines considered by the Progress out of Poverty group in Tanzania, converting the PPI score for each household into a poverty likelihood and averaging the poverty likelihood across all households in the sample yields an estimate for the percent of households in our sample below that poverty line.[44] We calculate that 53.1% of sample households are below 150% of the national poverty line in Tanzania and 68.3% are living on less than $2.00 per day in 2005 US dollars. Clearly, the sample is extremely resource constrained. The other asset indices are more difficult to interpret, but also show good balance across the two treatment arms.

Table 5.1 displays remarkable balance in characteristics across the two treatment arms. Of the 26 characteristics tested, just one difference is significant at the 5% level.[45] Even more reassuring, the normalized differences are extremely small in magnitude: none are above the 0.25 threshold and only two of 26 have a normalized difference above 0.10. Based on household demographic characteristics, physical

[44] When applying the PPI to a non-nationally representative group of households the poverty rate estimates are also potentially affected by an out of group bias.

[45] This difference is for whether the household head has some education, a characteristic for which there is little observed variation in the sample: 91.7% of control households and 94.3% of treatment households have a head with at least some education. This suggests that a significant difference in means between the treatment and control group is based on an imbalance in treatment assignment for a very small subset of people—only 198 households have a head with no education.

structure, amenities, and wealth, the randomization appears to have been extremely successful at selecting observably similar households.

Table 5.1: Demographics, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Household size	2,833	5.276	5.229	5.323	0.044	0.353
		(2.110)	(2.096)	(2.122)		
Female-Headed Household	2,833	0.129	0.133	0.125	-0.025	0.540
		(0.335)	(0.340)	(0.330)		
Age of Household Head	2,833	38.608	38.542	38.674	0.010	0.821
		(13.475)	(13.669)	(13.286)		
Household Head has some education	2,833	0.930	0.917	0.943	0.102	0.018
		(0.254)	(0.275)	(0.231)		
Household Head: Never Married	2,833	0.028	0.023	0.032	0.049	0.242
		(0.164)	(0.151)	(0.175)		
Household Head: Married, Monogamous	2,833	0.742	0.742	0.743	0.003	0.942
		(0.437)	(0.438)	(0.437)		
Household Head: Married, Polygamous	2,833	0.139	0.141	0.137	-0.011	0.813
		(0.346)	(0.348)	(0.344)		
Household Head's main activity is crop production	2,833	0.731	0.728	0.735	0.015	0.829
		(0.443)	(0.445)	(0.442)		
Age of the Primary Female	2,833	26.999	26.847	27.149	0.046	0.234
		(6.622)	(6.351)	(6.877)		
Primary Female has some education	2,833	0.936	0.926	0.946	0.082	0.115
		(0.245)	(0.262)	(0.226)		
Primary Female: Never Married	2,833	0.208	0.202	0.213	0.027	0.559
		(0.406)	(0.402)	(0.409)		
Primary Female: Married, Monogamous	2,833	0.677	0.679	0.675	-0.008	0.838
		(0.468)	(0.467)	(0.469)		
Primary Female: Married, Polygamous	2,833	0.089	0.090	0.088	-0.008	0.873
		(0.284)	(0.286)	(0.283)		
Order of wife - first	281	0.306	0.313	0.299	-0.029	0.768
		(0.462)	(0.465)	(0.460)		
Order of wife - second	281	0.598	0.611	0.584	-0.055	0.628
		(0.491)	(0.489)	(0.495)		
Order of wife - third	281	0.075	0.069	0.080	0.041	0.741
		(0.263)	(0.255)	(0.273)		
Order of wife - fourth	281	0.021	0.007	0.036	0.203	0.124
		(0.145)	(0.083)	(0.188)		
Duration of primary female living with most recent husband/partner (months)	2,099	88.763	86.479	91.019	0.062	0.204
		(73.083)	(72.147)	(73.962)		
Primary Female's main activity is crop production	2,833	0.768	0.762	0.774	0.029	0.660
		(0.422)	(0.426)	(0.419)		
Female focus child (under 12 months)	1,928	0.512	0.531	0.494	-0.073	0.107
		(0.500)	(0.499)	(0.500)		
Percent of school-aged children going to school	2,629	49.877	49.703	50.047	0.009	0.832
		(40.054)	(39.879)	(40.240)		
Total PPI score	2,833	39.337	39.478	39.198	-0.020	0.759

	N	All	Control	Treatment	Normalized Difference	P-value
		(14.146)	(14.263)	(14.033)		
Consumer Durable Asset Index	2,833	0.000	-0.013	0.013	0.012	0.849
		(2.065)	(2.143)	(1.985)		
Household Production Asset Index	2,833	-0.000	-0.038	0.038	0.053	0.234
		(1.428)	(1.468)	(1.387)		
Household Livestock Asset Index	2,833	-0.000	0.066	-0.065	-0.095	0.078
		(1.388)	(1.813)	(0.763)		
Household Total Asset Index	2,833	-0.000	-0.018	0.018	0.016	0.785
		(2.263)	(2.350)	(2.176)		
Household has access to improved sanitation	2,833	0.757	0.750	0.763	0.031	0.522
		(0.429)	(0.433)	(0.425)		
Household has access to improved drinking water sources	2,833	0.824	0.842	0.805	-0.096	0.297
		(0.381)	**(0.365)**	**(0.396)**		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

5.3 Primary Outcomes and Child Anthropometry

Table 5.2 presents the same set of observable characteristics for the primary outcomes and additional measures of child anthropometry as recorded during household survey. An adult in each of the households with a child under twelve months of age (1,885 across the two treatment groups) was asked for their consent to measure the height and weight of that child. In cases where the household had been identified in the CLE as having a pregnant woman but by the time the household survey was conducted the pregnant woman had already given birth, enumerators asked for permission to measure anthropometry for the recently born child (36 such children were successfully measured). The height and weight measurements were taken by enumerators who had been trained extensively as anthropometry specialists during the baseline enumerator trainings, and recorded by the household survey enumerator directly into the CAPI programme. These measurements were then immediately converted into Z-scores and malnutrition indicators (stunted, wasted, underweight, acutely malnourished) using the already recorded birth month and sex of the child being measured and following the procedures indicated by the World Health Organization child growth standards (WHO, 2006). For variables that have been converted into Z-scores (HAZ, WAZ, WHZ), values should be interpreted as standard deviations of the reference distribution of healthy children from five countries around the world. In Table 5.2, we show summary statistics for height (in cm) and weight (in kg); height-for-age Z-scores (HAZ), weight-for-age Z-scores (WAZ), and weight-for-height Z-scores (WHZ); indicators for whether the child is stunted (a HAZ below -2), the child is underweight (a WAZ below -2), and the child is wasted (a WHZ below -2); and indicators for whether the child is moderately acutely malnourished (WHZ between -2 and -3) or severely acutely malnourished (WHZ below -3). In addition to the summary measures contained in Table 5.2, we also show histograms of the distribution of HAZ, WHZ, and WAZ for the full sample in Figure 5.1, Figure 5.3, and Figure 5.5 and density plots of the distribution of HAZ, WHZ, and WAZ by treatment group in Figure 5.2, Figure 5.4, and Figure 5.6.

Table 5.2: Primary Outcomes, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Child age in months (under 12 months)	1,928	5.521	5.569	5.474	-0.028	0.541
		(3.421)	(3.450)	(3.393)		
Male child (under 12 months)	1,928	0.488	0.469	0.506	0.073	0.107
		(0.500)	(0.499)	(0.500)		
Child height	1,856	61.845	61.769	61.921	0.024	0.620

	N	All	Control	Treatment	Normalized Difference	P-value
		(6.306)	(6.319)	(6.296)		
Child weight	1,857	6.595	6.563	6.627	0.039	0.404
		(1.653)	(1.664)	(1.642)		
Child height-for-age Z-score	1,852	-1.425	-1.443	-1.407	0.030	0.613
		(1.190)	(1.225)	(1.156)		
Child weight-for-age Z-score	1,855	-0.670	-0.683	-0.656	0.023	0.649
		(1.182)	(1.227)	(1.137)		
Child weight-for-height Z-score	1,851	0.450	0.443	0.457	0.011	0.825
		(1.209)	(1.228)	(1.190)		
Child stunted	1,852	0.296	0.301	0.290	-0.024	0.656
		(0.457)	(0.459)	(0.454)		
Child underweight	1,855	0.121	0.133	0.110	-0.069	0.144
		(0.327)	(0.339)	(0.313)		
Child wasted	1,851	0.020	0.027	0.013	-0.103	0.028
		(0.140)	(0.163)	(0.113)		
Moderate acute malnutrition	1,851	0.016	0.021	0.011	-0.080	0.081
		(0.124)	(0.142)	(0.103)		
Severe acute malnutrition	1,851	0.004	0.007	0.002	-0.067	0.142
		(0.066)	(0.081)	(0.046)		
Children born in the last 24 months who were put to the breast within one hour	2,173	0.764	0.751	0.777	0.062	0.168
		(0.424)	(0.433)	(0.416)		
Infants 0-5 months of age who are fed exclusively with breast milk	1,129	0.013	0.005	0.021	0.136	0.042
		(0.115)	(0.073)	(0.143)		
Children 12-15 months of age who are fed breast milk	102	0.833	0.902	0.765	-0.371	0.069
		(0.375)	(0.300)	(0.428)		
Infants 6-8 months of age who receive solid, semi-solid or soft foods	461	0.696	0.691	0.701	0.021	0.829
		(0.460)	(0.463)	(0.459)		
Children 6-23 months of age who consume 4 or more food groups	1,038	0.214	0.205	0.222	0.040	0.570
		(0.410)	(0.404)	(0.416)		
Number of food groups (of 7) children 6-23 months of age consume	1,038	2.366	2.329	2.402	0.052	0.482
		(1.406)	(1.360)	(1.451)		
Children 6-23 months of age who meet the minimum meal frequency	1,044	0.455	0.457	0.453	-0.009	0.901
		(0.498)	(0.499)	(0.498)		
Women's Dietary Diversity Score (1-9)	2,833	3.529	3.481	3.577	0.086	0.168
		(1.112)	(1.103)	(1.120)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

On average, children are 5.5 months of age at the time of the survey—5.6 in control villages and 5.5 in treatment villages—and 48.8% are male—46.9% in control villages and 50.6% in treatment villages. Relative to the reference distribution of healthy children, the sample children are substantially shorter given their age and sex: the mean HAZ score is -1.4 in both treatment and control villages. While the distribution of weight appears to be closer to that of the reference group (mean WAZ of -0.670 overall, -0.683 in control villages, and -0.656 in treatment villages), it still lags behind the mean weight for the children used to construct the reference growth standards. Because height for children in the sample was considerably lower than weight relative to their respective reference distributions, WHZ scores are, on average, positive and wasting is extremely uncommon: mean WHZ is 0.450 overall, 0.443 in control villages, and 0.457 in treatment villages while only 2.0%, 2.7%, and 1.3% of children are wasted overall, in the control villages, and in treatment villages. Unsurprisingly, moderate and severe acute malnutrition is even less common in the sample with just 0.4% of children severely acutely malnourished and just another 1% of children

moderately acutely malnourished. Stunting, on the other hand, is quite common among the measured children: 29.6% of children overall are stunted, 30.1% in control villages, and 29.0% in treatment villages. Given the now extensive literature linking early childhood stunting to worse later-in-life outcomes (Black, 2013), the observed low levels of height are concerning and reflective of poor early childhood nutrition.

For several of the primary outcome measures, we can compare the means from the baseline data to means for the region of Iringa from the 2015-2016 DHS (Boyle et al., 2017). This comparison is useful for understanding whether the eligibility restrictions made for this study—namely, that households have at least one member who is literate in Swahili and own a mobile phone—affect the levels of the primary outcomes of interest. However, we caution that there are several important caveats to this comparison. First, because of our focus on children during the first 1,000 days, all the children measured in our sample are under twelve months of age. Region-level means are available in the DHS data for children under the age of five, implying that there will be a substantial difference in the age distribution of measured children. These differences in age are relevant determinants of HAZ and stunting (Victoria et al., 2010), but they also may indirectly affect the other primary outcomes if there are age-specific feeding patterns in the study region or if we are more likely to select households that differ socioeconomically from the general population of Iringa because of the child age restriction. Nevertheless, the contrast can still provide potentially useful context for the evaluation.

Mean HAZ for children in Iringa region in the 2015-2016 DHS is -1.8 and 41.6 of measured children under five years of age are stunted. As mentioned above, for the children in our sample, mean HAZ is -1.4, with 29.6% of children classified as stunted. This suggests that our sample may be slightly better off in terms of child nutrition. However, as mentioned earlier, a comparison of child growth patterns in 54 countries with the WHO child growth standards found that HAZ scores decline relative to the standard until children reach around 24 months of age. The difference in age between the children in our data and those in the DHS data can potentially explain the gap in mean HAZ and stunting between the two samples.

DHS levels for the three other IYCF practices observable in both surveys, the percent of children who started breastfeeding within one hour of birth, the percent of children 6-23 months fed from four or more food groups, and the percent of children 6-23 months that meet the minimum meal frequency, are somewhat different from those found in our data. For example, 58.2% of children in Iringa in the DHS began breastfeeding within one hour of birth, 27.5% of children 6-23 months were fed from four or more food groups during the day and night preceding the survey, and 20.2% met the minimum meal frequency requirements given their age and breastfeeding status. In our baseline, we find that 76.4% of children born with in the past two years were breastfed within one hour of birth, 21.4% of children 6-23 months were fed from four or more food groups during the day and night preceding the survey, and 45.5% met the minimum meal frequency requirements. While the data are not directly comparable—our data collection excludes all urban households in the region and all households that do not have a child under twelve months of age or a pregnant woman—if anything, these differences indicate that the children in our sample may be slightly advantaged relative to children from households surveyed during the 2015-2016 DHS in Iringa. Conversely, we identify a rate of exclusive breastfeeding for children 0-5 months of age of just 1.3%. The DHS data do not ask mothers to report on the exclusive breastfeeding status of all children 0-5 months of age. However, mothers do report on when children were first fed anything other than breastmilk. Using this information and limiting the sample to children 0-5 months of age at the time of the DHS indicates that just 4.3% of children across Tanzania within that age range have never been fed anything other than breastmilk. Among households in Iringa, this figure rises to 8.9%, though there are just 16 children within this age range in the 2016 and our estimate of 1.3% is not statistically significantly different from the analogous DHS estimate.

As with Table 5.1, Table 5.2 indicates that the two treatment groups are well balanced in terms of baseline measures of the anthropometry outcomes. The largest observed normalized difference is -0.103 (for the Child is wasted indicator), still well below the 0.25 cut-off. Similarly, only one of the twelve tests of the null of no difference between the mean value in treatment villages and in control villages has a p-value below 0.05 (for wasting prevalence); this, again, is roughly what we should expect to observe by chance.

Figure 5.1, Figure 5.3, and Figure 5.5 display histograms for the full sample distribution of HAZ, WAZ, and WHZ. The distributions are clearly unimodal and appear to approximate a normal distribution somewhat well. The density plots in Figure 5.2, Figure 5.4, and Figure 5.6, which show the distribution of HAZ, WAZ, and WHZ separately by treatment group, uncover no notable differences in the distributions by treatment status. We also conduct Kolmogorov-Smirnov tests of the equality of distributions for each of the three indicators. In effect, Kolmogorov-Smirnov tests check for differences in the empirical cumulative density functions (the probability that a variable is less than or equal to a particular value) by treatment status. A failure to reject the null of no difference, is indicative of the observed distributions for an indicator being close to one another among observations in the treatment group and observations in the control group. The p-values from the Kolmogorov-Smirnov tests are 0.63, 0.67, and 0.76 for HAZ, WAZ, and WHZ, respectively, suggesting that there are no important differences in the distributions of anthropometry across the two treatment groups.

Figure 3: Distribution of Children's HAZ

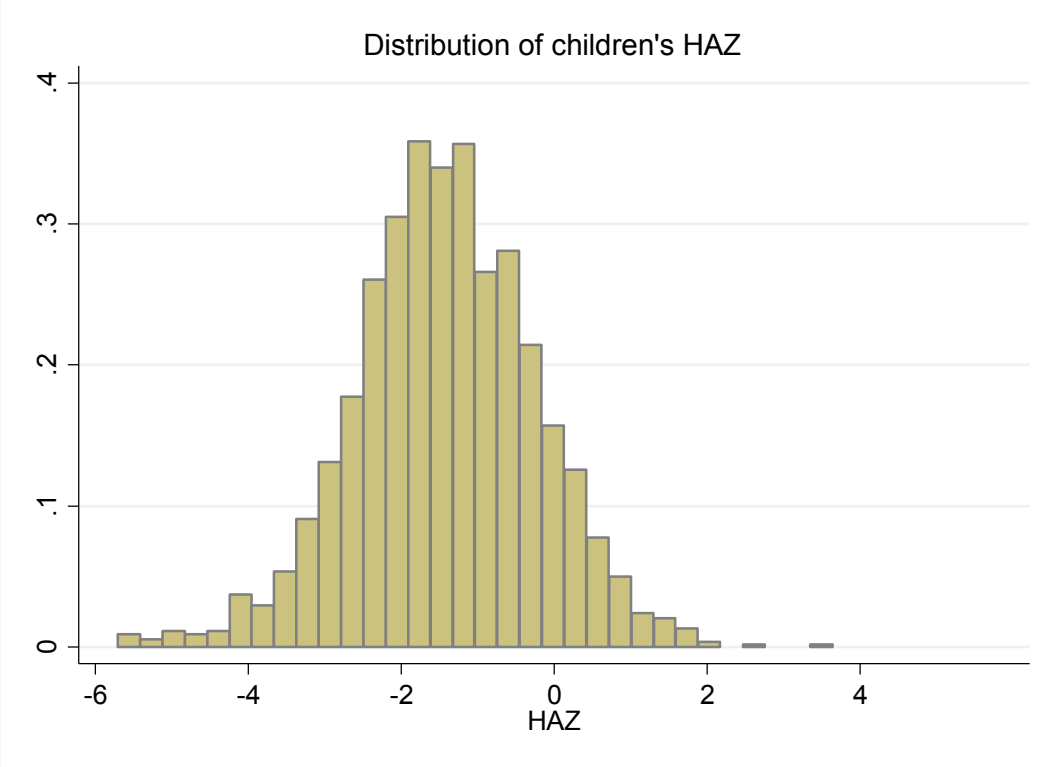


Figure 4: Distribution of children's HAZ, by mNutrition Treatment Status

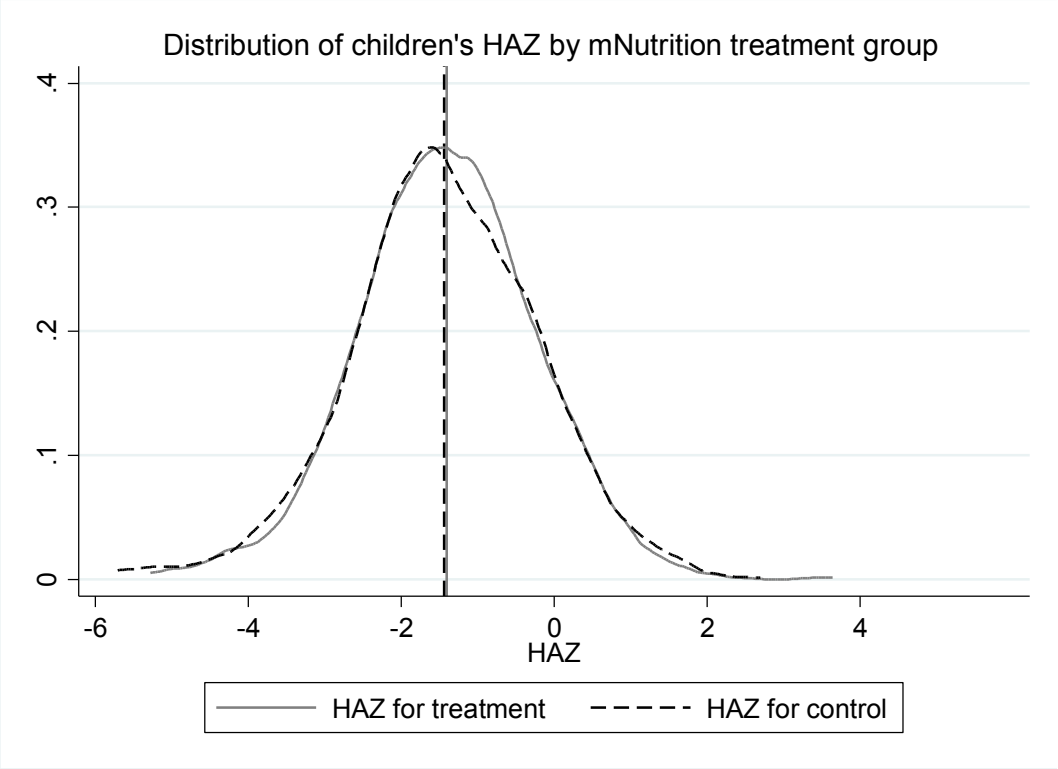


Figure 5: Distribution of children's WHZ

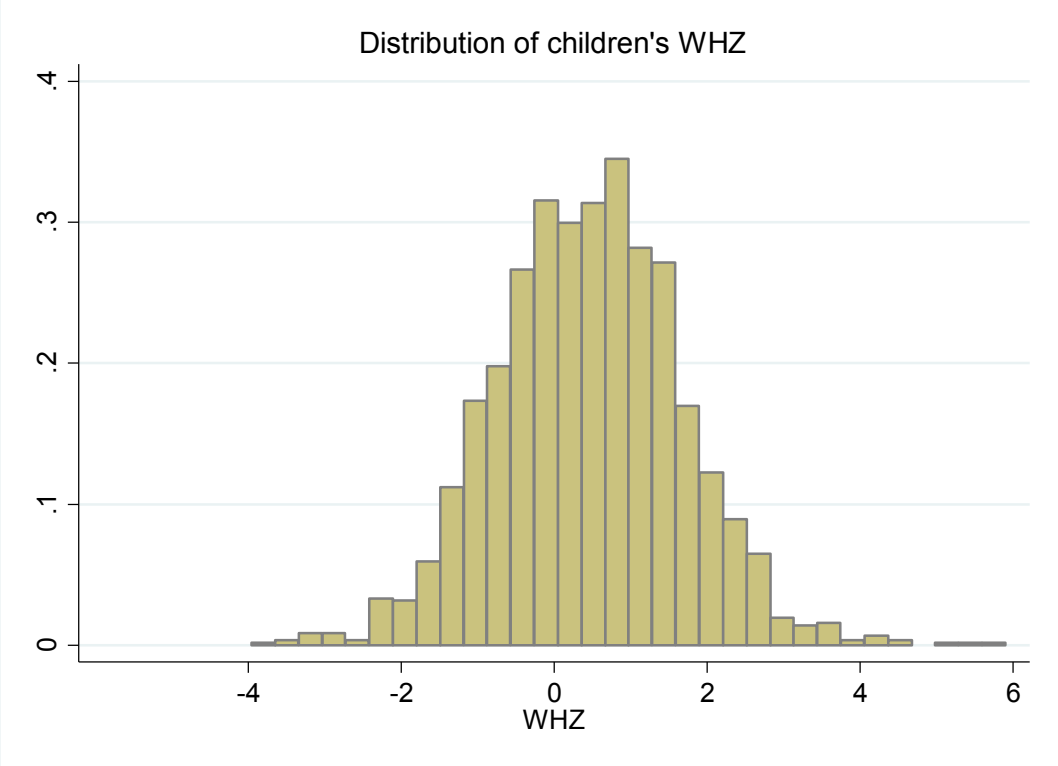


Figure 6: Distribution of children's WHZ, by mNutrition Treatment Status

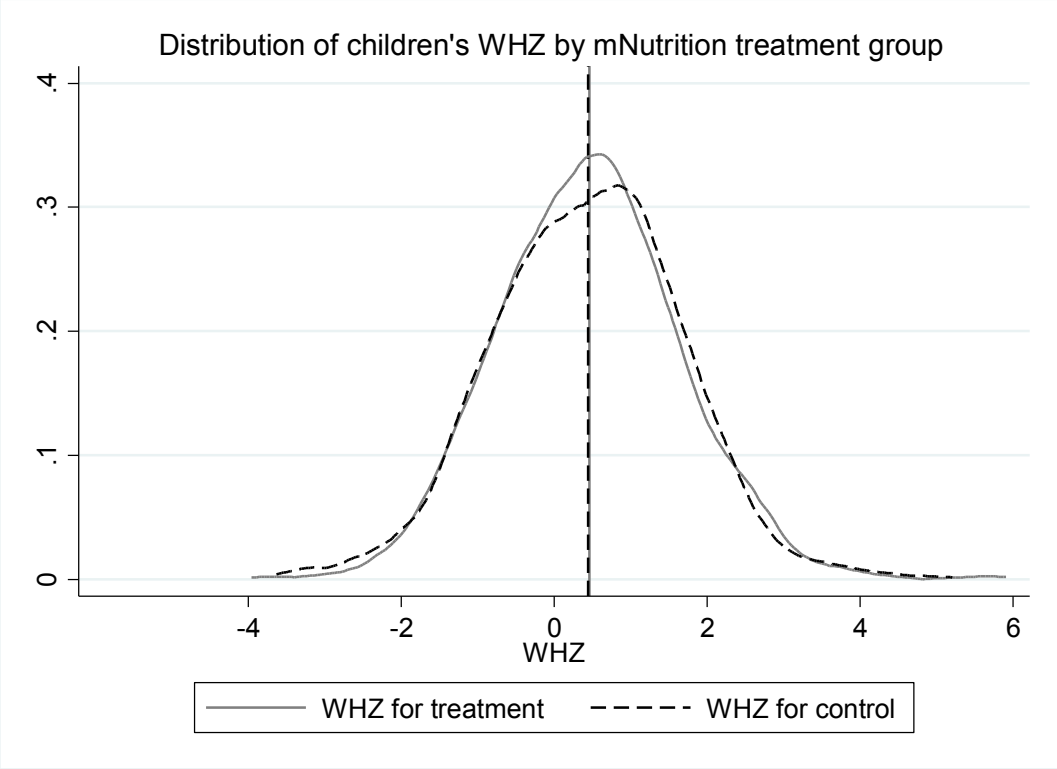


Figure 7: Distribution of children's WAZ

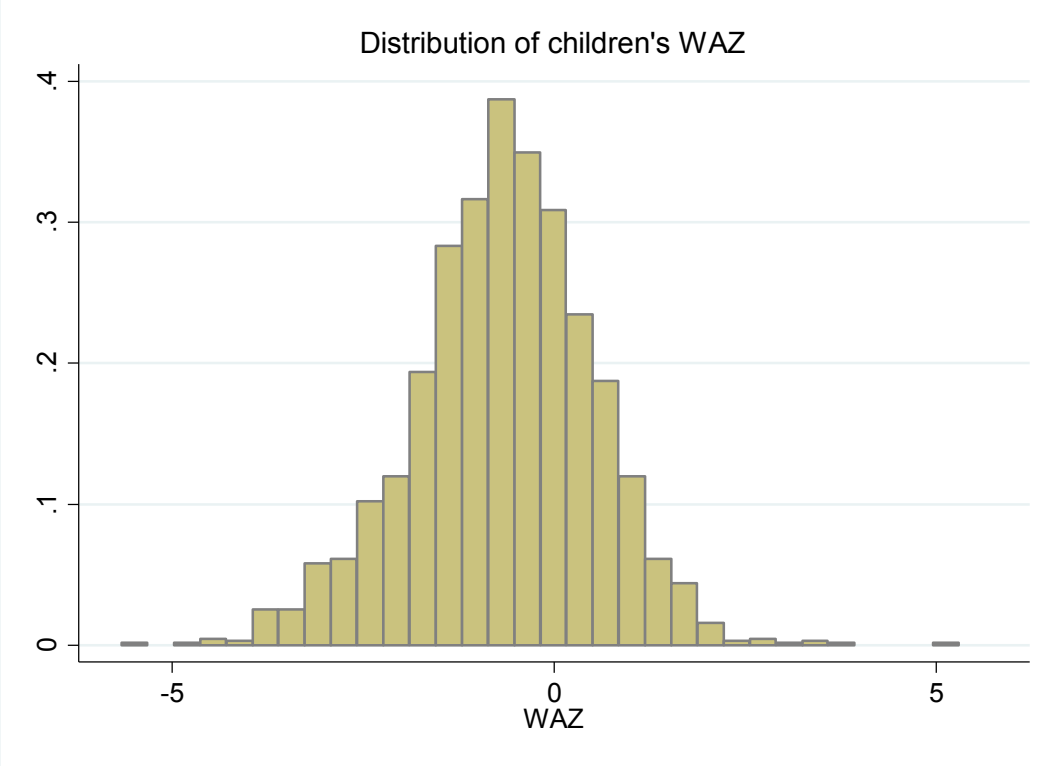
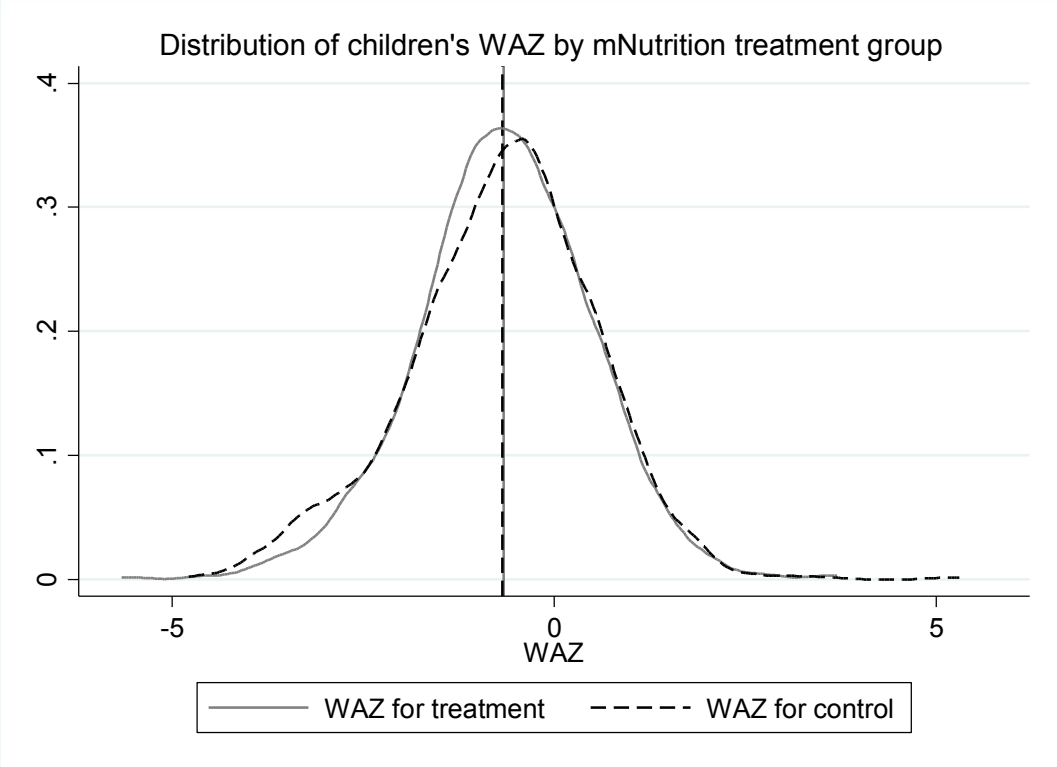


Figure 8: Distribution of children's WAZ, by mNutrition Treatment Status



Infant and Young Child Feeding Practices (IYCF) are based on mother reports on past practices for up to two children: any child under twelve months of age and the next youngest child (up to age five) of the mother of the child under twelve months or the pregnant woman. As some of the IYCF practices are highly dependent on child age—the mother is not asked to provide information on whether a five-month-old is receiving food from four or more food groups—the values for those variables include only the responses that pertain to children that are at least as old as the lower end of the relevant age range. Each of the IYCF practices is generated as an indicator equal to one if the mother reports that she followed the recommended behaviour (e.g. fed her child 6-23 months of age solid, semi-solid, or soft food) and zero otherwise. Children who are not old enough for the question to be relevant, are coded as having a missing value for that question.

The six core IYCF practices indicators that we consider are whether the child was put to the breast within one hour of birth, whether the child was exclusively breastfed during the first six months, whether the child was still fed breast milk between twelve and fifteen months of age, whether the child received foods from four or more food groups during the day preceding the survey, and whether the child met minimum meal frequency standards in the past 24 hours given their age and breastfeeding status.[46] We also show the raw number of food groups consumed by the child during the 24 hours preceding the survey.

Table 5.2 indicates that there is important heterogeneity in IYCF practices across the various indicators. For example, exclusive breastfeeding for the first six months, minimum meal frequency, and children's dietary diversity measures suggest that IYCF practices for children in the sample are not adequate. Only 1.3% of children were exclusively breastfed for the first six months (0.5% in control villages and 2.1% in treatment villages), 21.4% of children 6-23 months of age receive food from four or more food groups (20.5% in control villages and 22.2% in treatment villages), and only 45.5% of children satisfy minimum meal frequency given their age and breastfeeding status (45.7% in control villages and 45.3% in treatment villages). Conversely, households seem to perform much better when it comes to early initiation of breastfeeding, 76.4% of children were put to the breast within one hour of birth (75.1% in the control group and 77.7% in the treatment group), 69.6% of infants 6-8 months of age received solid, semi-solid, or soft

[46] The minimum meal frequency (during the past 24 hours) for non-breastfed children is four for children 6-23 months of age. For breastfed children: the minimum meal frequency is two for children 6-8 months of age and three for children 9-23 months of age.

foods (69.1% in control villages and 70.1% in treatment villages), and 83.3% of children are still fed breast milk between the ages of 12 and 15 months (90.2% in control villages and 76.5% in treatment villages). Still, there appears to be room for improvement for all the IYCF practices indicators, albeit more for some of the indicators (exclusive breastfeeding for six months; child dietary diversity; minimum meal frequency), than others (early initiation of breastfeeding; consumption of solid, semi-solid, or soft foods for children 6-8 months of age; continued breastfeeding for children 12-15 months of age).

As with the anthropometry measures, IYCF practices are well balanced across the two treatment arms. Only one normalized difference is above the 0.25 threshold—the normalized difference for children 12-15 months still being fed breast milk is -0.37, suggesting that children in treatment villages are less likely to still be fed breast milk between those ages. However, we would encourage caution when interpreting this difference, as there are only 102 children in total who fall within this age range and the p-value from the t-test of equality of means between the treatment and control groups fails to reject the null hypothesis at the 5% level. The remainder of the normalized differences are well below the 0.25 cut-off, and only one difference is significant at the 5% level: infants were slightly more likely to be exclusively breastfed for the first six months in treatment villages (p-value 0.042). While we would prefer to observe no normalized differences greater than 0.25 in absolute value and no p-values below 0.05, it is likely given the number of tests that we are conducting that there will be some imbalance eventually. Further, it is reassuring that the two differences in this table do not favour one treatment group over the other: treatment households are more likely to exclusively feed breast milk to children for the first six months but less likely to continue breastfeeding children between the ages of 12-15 months. It therefore seems unlikely that these differences are a symptom of more problematic differences in unobservable characteristics between the two groups.

Women's dietary diversity is assessed using the Women's Dietary Diversity Score (Kennedy et al., 2013). The Women's Dietary Diversity Score (WDDS) is intended to reflect the nutritional quality of a woman's diet by measuring the likelihood that the woman is achieving micronutrient adequacy. The score is a count of the number of food groups, out of a maximum of nine, from which the woman consumed during the 24-hour period preceding the survey.[47]

Women's dietary diversity scores are also low among the primary females in the sample. The mean number of food groups consumed—from starchy staples; dark green leafy vegetables; other vitamin A rich fruits and vegetables; other fruits and vegetables; organ meat; meat and fish; eggs; legumes, nuts, and seeds; milk and milk products—is just 3.529 overall, 3.481 in control villages, and 3.577 in treatment villages. This suggests that most women would not achieve the minimum level of women's dietary diversity (MDD-W), a measure that is defined using a slightly different set of food groups but which requires that women consume from at least five out of the ten possible food groups in the previous 24 hours (FAO, 2016). The treatment groups are also well balanced in WDDS—the normalized difference is just 0.09 and the p-value from the t-test of no difference between the treatment and control group is 0.17.

The preceding tables and figures suggest that there are considerable gaps in nutrition outcomes and behaviours for the households included in the sample. HAZ for children, WDDS, and four of the six IYCF practices investigated are below desired levels. While these gaps are certainly not sufficient to indicate that the mNutrition intervention will be successful, they do at least indicate that there are plausible pathways through which the programme could improve outcomes; we explore this theory further in the next section on IYCF Knowledge and Beliefs. Consistent with the results from Table 5.2, the randomization appears to have been successful in balancing the baseline values of the primary outcomes across treatment and

[47] We elected to use the WDDS instead of the Minimum Dietary Diversity for Women indicator (MDD-W) because we expect that the more continuous nature of the WDDS will improve the statistical power we have to detect differences between the treatment and control groups. That said, we also plan to investigate whether the mNutrition programme impacts MDD-W. Additionally, we use the list-based method for eliciting food group consumption at the household and individual levels (FAO, 2016). We do so because CAPI programming of this method was more straightforward. To ensure that enumerators were able to completely characterize the respondents' diets during the previous day all enumerators were carefully trained how to ask follow-up questions about food additions that may have been used (e.g. milk in tea) but not immediately listed by the respondent.

control households: normalized differences are well below the 0.25 threshold and there are few statistically significant differences in means.

5.4 Secondary Outcomes: IYCF Knowledge and Beliefs

Table 5.3 displays indicators of the primary female and primary male Infant and Young Child Feeding knowledge and beliefs. We display the percent of correct answers that the respondent gave as a summary measure of IYCF knowledge and beliefs, constructed from the eleven questions contained in Module H parts 1 and 2. The eleven questions regarding IYCN knowledge and beliefs covered breastfeeding, complementary feeding, and other health and nutrition topics.

Table 5.3: Secondary Outcomes, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Infant and Young Child Knowledge (female)						
Percentage of correct answers	2,833	55.948	55.315	56.570	0.071	0.164
		(17.749)	(18.222)	(17.255)		
Infant and Young Child Knowledge (male)						
Percentage of correct answers	1,508	48.463	48.951	47.991	-0.040	0.502
		(24.165)	(24.051)	(24.281)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 5.3 separately displays the measures of IYCF Knowledge and Beliefs for the primary females and primary males who were interviewed. While women have substantially more knowledge (or more accurate beliefs) about IYCF practices—on average, women correctly answered 55.9% of the IYCF questions correctly compared to just 48.5% among males—there are clearly substantial gaps in knowledge for most adults in the sample. This provides more evidence on how the informational treatment offered by the mNutrition programme may be able to affect nutrition outcomes: by improving the nutrition related knowledge and beliefs of adults who may then select to engage in better IYCF or dietary practices. Further, the lower levels of nutrition knowledge among the men who participated in the baseline survey indicates that providing the information directly to men in the treatment households may have more scope to change beliefs, and potentially household-level practices (see research question 5).

As with the two prior tables, the randomization appears to have balanced baseline levels of IYCF knowledge and beliefs across the two treatment groups: the largest normalized difference is well below 0.10 and neither of the differences are significant at the 5% level.

5.5 Sources of Nutrition and Health Information

To explore where sample households get their information on nutrition and health, we use two modules of the baseline household survey. The first asks the primary female respondents whether they consider a range of different possible organizations and people to be sources for health and nutrition information and, subsequently, which sources are the most important for them. The next uses a five point Likert scale, ranging from “Strongly Disagree” to “Strongly Agree” to measure the level of trust that respondents would have in health and nutrition information from different sources.

Understanding where households currently get information on health and nutrition related topics as well as how much they would trust information coming from different sources—in particular, from SMS messages—is critical to speculate whether the mNutrition intervention is likely to have any impact on behaviours. If individuals indicate a willingness to trust health and nutrition information that they receive on their mobile phones, it suggests that the sample may be more likely to change their beliefs and behaviours in response to a mobile based intervention. Conversely, if individuals in the sample report not being willing to trust

information they receive through SMS messages, then there may be little hope of the intervention changing beliefs, behaviours, and outcomes.

Table 5.4: Sources of information and trust likelihood on health and nutrition, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
The following is a source of information:						
Spouse	2,173	0.523	0.499	0.546	0.094	0.107
		(0.500)	(0.500)	(0.498)		
Family	2,833	0.405	0.383	0.427	0.090	0.091
		(0.491)	(0.486)	(0.495)		
Friends/neighbours	2,833	0.458	0.431	0.485	0.108	0.039
		(0.498)	(0.495)	(0.500)		
Automated text messages	2,833	0.274	0.256	0.291	0.080	0.158
		(0.446)	(0.436)	(0.455)		
Government health workers	2,833	0.968	0.967	0.969	0.011	0.807
		(0.175)	(0.178)	(0.173)		
Non-government health facilities	2,833	0.449	0.443	0.456	0.026	0.611
		(0.498)	(0.497)	(0.498)		
TV/Radio/Posters	2,833	0.690	0.680	0.700	0.043	0.333
		(0.463)	(0.467)	(0.459)		
Traditional health workers	2,833	0.065	0.063	0.067	0.019	0.664
		(0.246)	(0.242)	(0.251)		
Non-government health workers	2,833	0.387	0.381	0.393	0.023	0.615
		(0.487)	(0.486)	(0.489)		
Community health workers	2,833	0.579	0.566	0.591	0.051	0.319
		(0.494)	(0.496)	(0.492)		
*The following is the **primary source** of information:*						
Spouse	2,833	0.143	0.136	0.151	0.042	0.345
		(0.350)	(0.343)	(0.358)		
Family	2,833	0.067	0.064	0.071	0.027	0.512
		(0.251)	(0.245)	(0.256)		
Friends/neighbors	2,833	0.030	0.021	0.038	0.097	0.014
		(0.170)	(0.145)	(0.191)		
Automated text messages	2,833	0.020	0.017	0.023	0.043	0.280
		(0.140)	(0.130)	(0.150)		
Government health workers	2,833	0.802	0.809	0.796	-0.033	0.501
		(0.399)	(0.394)	(0.403)		
Non-government health facilities	2,833	0.131	0.135	0.127	-0.023	0.693
		(0.337)	(0.341)	(0.333)		
TV/Radio/Posters	2,833	0.119	0.120	0.118	-0.008	0.840
		(0.324)	(0.325)	(0.322)		
Traditional health workers	2,833	0.002	0.002	0.001	-0.017	0.638
		(0.042)	(0.046)	(0.037)		
Non-government health workers	2,833	0.040	0.036	0.044	0.044	0.278
		(0.196)	(0.185)	(0.205)		
Community health workers	2,833	0.098	0.099	0.097	-0.005	0.914
		(0.298)	(0.299)	(0.297)		

	N	All	Control	Treatment	Normalized Difference	P-value
*Agree they can **trust** the following source of information:*						
Spouse	2,173	0.844	0.831	0.856	0.068	0.217
		(0.363)	(0.375)	(0.352)		
Family	2,833	0.694	0.690	0.699	0.020	0.746
		(0.461)	(0.463)	(0.459)		
Friends/neighbors	2,833	0.582	0.572	0.593	0.044	0.417
		(0.493)	(0.495)	(0.491)		
Automated text messages	2,833	0.820	0.801	0.838	0.096	0.069
		(0.384)	(0.399)	(0.368)		
Government health workers	2,833	0.984	0.986	0.982	-0.032	0.376
		(0.124)	(0.116)	(0.131)		
Private clinic/hospital	2,833	0.883	0.889	0.877	-0.038	0.443
		(0.322)	(0.314)	(0.329)		
TV/radio/posters	2,833	0.872	0.869	0.874	0.015	0.718
		(0.335)	(0.337)	(0.332)		
Traditional health worker	2,833	0.098	0.090	0.106	0.054	0.198
		(0.297)	(0.286)	(0.308)		
Non-government health worker	2,833	0.790	0.797	0.784	-0.033	0.493
		(0.407)	(0.402)	(0.412)		
Community health worker	2,833	0.911	0.907	0.915	0.027	0.487
		(0.285)	(0.291)	(0.280)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 5.4 presents the baseline values for key information sources and trust in health and nutrition information from different sources. Women report government health workers as the most common source of information for health and nutrition information: 96.8% of women report government health workers as an information source. The next most common source of information is TV/Radio/Posters (69.0%), followed by community health workers (57.9%). At the other end of the distribution, only 6.5% of women report that traditional health workers are a source of nutrition and health information and just over a quarter (27.4%) report that automated text messages are a source of information. Consistent with this, 80.2% of women report that government health workers are their primary source of nutrition and health information, followed by non-governmental health facilities (13.1%), and TV/Radio/Posters (11.9%).

Despite just 27.4% of women currently receiving information through automated text messages, an extremely high fraction (82.0%) of women report that they agree with the statement that “If I were to receive any breastfeeding, complementary feeding, or general health information from automated SMS messages from an NGO, private organization, religious, or voluntary organization or the government, I would feel confident and trust it completely.” In fact, this figure is nearly in line with the fraction of women who would trust information from their spouse (84.4%), from a private clinic/hospital (88.3%), or from TV/Radio/posters (87.2%), and it is similar to the fraction of women who report that they would trust information from a non-governmental health worker (79.0%), from friends/neighbours (58.2%), or from their family (69.4%); trust in information from government health workers (98.4%) is by far the highest. Still, the self-reported hypothetical trust in health and nutrition information received via automated text messages is promising. It suggests that the women in the sample may be responsive to information received through the mNutrition programme. Taken at face value, the results in Table 5.4 indicate that they may even be more responsive to the mNutrition messaging than they would be to information from non-governmental health workers, friends and neighbours, and family members excluding their spouse.

None of the normalized differences in Table 5.4 cross the 0.25 threshold—the largest normalized difference is just 0.11, for whether friends/neighbours are a source of health and nutrition information. Similarly

reassuring, only two of the thirty tests of the null hypothesis of no difference in means between the treatment and control group result in p-values below 0.05.

5.6 Mobile Phone Access and Usage

Table 5.5 and Table 5.6 present baseline information on mobile phone access and usage as reported by the primary female and primary male, respectively. The data used to construct the indicators considered in this section are taken directly from responses given in Module F, Part 1a and Module F, Part 1b of the baseline household survey. Below we discuss the indicators separately by respondent sex, first for females and then for males, before briefly discussing gendered differences in the responses at the end of this sub-section.

5.6.1 Primary Females

Table 5.5: Mobile phone access and usage (female), by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Owns a mobile phone	2,833	0.589	0.576	0.602	0.054	0.269
		(0.492)	(0.494)	(0.490)		
Used a mobile phone in the last 14 days to						
make calls	2,833	0.709	0.694	0.723	0.065	0.151
		(0.454)	(0.461)	(0.447)		
receive calls	2,833	0.765	0.757	0.773	0.037	0.386
		(0.424)	(0.429)	(0.419)		
write text messages	2,833	0.486	0.475	0.497	0.045	0.346
		(0.500)	(0.500)	(0.500)		
receive text messages	2,833	0.624	0.617	0.632	0.030	0.529
		(0.484)	(0.486)	(0.483)		
send mobile money	2,833	0.090	0.094	0.086	-0.027	0.519
		(0.286)	(0.292)	(0.281)		
receive mobile money	2,833	0.152	0.150	0.154	0.011	0.816
		(0.359)	(0.357)	(0.361)		
use mobile internet	2,833	0.043	0.042	0.044	0.010	0.804
		(0.203)	(0.201)	(0.205)		
No mobile phone use in the last 14 days	2,833	0.156	0.159	0.154	-0.013	0.774
		(0.363)	(0.366)	(0.361)		
Mobile phone used often in the last 14 days	2,390	0.710	0.697	0.723	0.056	0.279
		(0.454)	(0.460)	(0.448)		
Used mobile phone to receive health advice	2,833	0.085	0.070	0.099	0.104	0.015
		(0.279)	(0.256)	(0.299)		
Received automatic text messages (nutrition information) in past 2 years	2,833	0.072	0.063	0.082	0.075	0.106
		(0.259)	(0.242)	(0.274)		
Amount spent on airtime on all phones in an average month (Tz shillings)	2,833	2,450.124	2,356.726	2,542.017	0.041	0.368
		(4,471.412)	(4,305.505)	(4,628.518)		
Charges phone at home	2,833	0.455	0.452	0.457	0.011	0.853
		(0.498)	(0.498)	(0.498)		
Takes less than 30 minutes to get to the nearest place to charge their phone	1,545	0.968	0.968	0.969	0.009	0.888
		(0.175)	(0.177)	(0.173)		
	1,545	988.982	921.434	1,056.094	0.081	0.143

	N	All	Control	Treatment	Normalized Difference	P-value
Amount spent on charging phone in an average month (Tz shillings)		(1,654.640)	(1,457.481)	(1,828.098)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Mobile phone ownership among the primary females is high: 58.9% of the primary females in the overall sample own a mobile phone, 57.6% in control villages, and 60.2% in treatment villages. While making and receiving phone calls is the most common way in which women in the sample used a mobile phone in the past 14 days—70.9% used a mobile to make a phone call and 76.5% to receive a phone call—sending and receiving text messages is still common in the sample: 62.4% of women used a mobile phone to receive a text message and 48.6% used a mobile phone to send a text message in the past 14 days. A non-trivial fraction of the women in the sample also used their phone to send (9.0%) and receive (15.2%) mobile money. Mobile internet use, however, is uncommon (4.3%). Consistent with these behaviours, just 15.6% of the women in the sample reported not having used a mobile phone at all during the last 14 days. Use of mobile phones to receive health advice is lower than reported in the previous section (8.5% as compared to 27.4%), likely as a result of the shorter recall period—just two weeks—and just 7.2% of women report having received nutrition information through text messages during the past two years.

Just shy of half (45.5%) of women report that they typically charge their mobile phone at home, and nearly the entire sample (96.8%) charges their phone at a place that is less than 30 minutes away from their home. This suggests that women have access to, and frequently own their own, mobile phones, and there do not appear to be large physical or spatial barriers to regular use of mobile phones.

Women report spending, on average, 2,450.1 Tanzanian Shillings on airtime per month and another 988.4 Shillings on charging their phones per month. In total, this suggests that women spend on average 3,439.1 on airtime and charging per month, or around $1.5 US.

Differences between the treatment and control group in primary female mobile phone use are small in magnitude: the largest normalized difference is 0.10 (for whether women used their mobile phone to receive health advice). Most of the remaining normalized differences are below 0.05, and only one of the t-tests indicates that there is a statistically significant difference in means between the treatment and control households at the 5% level (again, whether women used their mobile phone to receive health advice). Given the number of tests (16 for this section), we should expect to find between 0 and 1 statistically significant differences simply by chance. Table 5.5 therefore continues to suggest that the randomization was successful at balancing observable characteristics across the treatment arms.

5.6.2 Primary Males

Table 5.6: Mobile phone access and usage (male), by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Owns a mobile phone	1,404	0.907	0.907	0.907	0.000	0.997
		(0.290)	(0.290)	(0.290)		
Used a mobile phone in the last 14 days to						
make calls	1,404	0.896	0.890	0.902	0.039	0.507
		(0.305)	(0.313)	(0.298)		
receive calls	1,404	0.927	0.926	0.928	0.009	0.883
		(0.260)	(0.262)	(0.258)		
write text messages	1,404	0.645	0.630	0.661	0.065	0.334
		(0.479)	(0.483)	(0.474)		
receive text messages	1,404	0.835	0.842	0.827	-0.040	0.494
		(0.372)	(0.365)	(0.378)		
send mobile money	1,404	0.224	0.210	0.237	0.065	0.227
		(0.417)	(0.407)	(0.426)		
receive mobile money	1,404	0.278	0.278	0.278	-0.000	0.995
		(0.448)	(0.448)	(0.448)		
use mobile internet	1,404	0.077	0.071	0.083	0.044	0.449
		(0.267)	(0.257)	(0.276)		
No mobile phone use in the last 14 days	1,404	0.040	0.035	0.045	0.052	0.328
		(0.196)	(0.183)	(0.207)		
Mobile phone used often in the last 14 days	1,348	0.895	0.891	0.899	0.026	0.648
		(0.307)	(0.312)	(0.302)		
Used mobile phone to receive health advice	1,404	0.080	0.080	0.081	0.006	0.911
		(0.272)	(0.271)	(0.274)		
Received automatic text messages (nutrition information) in past 2 years	1,404	0.067	0.072	0.062	-0.043	0.457
		(0.250)	(0.259)	(0.241)		
Amount spent on airtime on all phones in an average month (Tz shillings)	1,404	7,941.880	7,936.324	7,947.264	0.001	0.984
		(8,581.123)	(8,411.974)	(8,747.834)		
Charges phone at home	1,404	0.474	0.472	0.475	0.007	0.918
		(0.499)	(0.500)	(0.500)		
Takes less than 30 minutes to get to the nearest place to charge their phone	739	0.970	0.964	0.976	0.068	0.413
		(0.170)	(0.186)	(0.153)		
Amount spent on charging phone in an average month (Tz shillings)	739	1,813.700	1,831.025	1,796.791	-0.014	0.867
		(2,515.213)	(2,647.253)	(2,382.744)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Of the primary males who were interviewed, 90.7% owned their own mobile phone in both the treatment and the control group. Perhaps unsurprisingly given the high level of mobile phone ownership, the primary males are also extremely likely to have used their mobile in the last 14 days to make (89.6%) or receive (92.7%) phone calls, or to write (64.5%) or receive (83.5%) text messages. Interestingly, while mobile ownership rates and the use of mobiles for phone calls are much higher among males than females, there appears to be less of a gap when it comes to receiving (83.5% for men and 62.4% for women) or writing (64.5% for men and 48.6% for women) text messages.

Of primary males, 22.4% used their mobile to send mobile money and 27.8% used their mobile to receive mobile money in the past 14 days. These figures represent roughly double the rates reported by the primary female sample. Mobile internet use is also approximately twice as likely for primary males (7.7%)

relative to primary females (4.3%). Finally, just 4.0% of primary males did not use a mobile phone during the past 14 days, suggesting nearly all of the male sample has regular access to a mobile phone.

Despite the increased ownership, access, and usage of mobile phones relative to the primary female sample, primary males are actually less likely to have used their mobile phone to receive health advice (8.0%) or to receive automatic text messages with nutrition information (6.7%), though the difference is not statistically significantly different from zero. Unfortunately, given the data we collected, it is impossible to identify whether this deficit relative to the primary females is driven by less interest in receiving health and nutrition information or some other economic, cultural, or social phenomenon.

Men spend considerably more on airtime and phone charging per month than their female counterparts: 7,941.9 Shillings and 1,813.7 Shillings, respectively. Across the two, men spend 9,755 Shillings (or approximately $4.4 US) per month, on average. This is nearly three times the amount that women spend on airtime and phone charging; this difference is likely a reflection of the higher mobile phone usage among men, and potentially also a function of men's increased control over resources—though the data do not directly speak to this latter possibility.

The largest normalized difference in Table 5.6 is 0.07, for the likelihood that men who do not primarily charge their mobile at home have to travel less than 30 minutes to get to the nearest place to charge their phone. Further, there are no statistically significant differences between the treatment and control group means for any of the indicators tested: the smallest p-value is 0.23, for whether individuals used their mobile phone in the past 14 days to send mobile money. As in previous tables, the evidence suggests that the treatment and control groups are extremely well balanced in observable characteristics—in this case, there are no important differences for any of the male mobile phone access and usage indicators.

5.7 General Child Health and Vaccination Compliance

Module D, Part 1 of the household survey collects information on general child health and vaccination histories for all children in the household under the age of five at the time of the interview. All of the answers are reported by the primary female, who, in most cases is the mother of all of the children under the age of five in the household.

Table 5.7: Child Health, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
General health rated good, very good, or excellent	4,019	0.821	0.812	0.830	0.045	0.224
		(0.383)	(0.391)	(0.376)		
Was ill with a fever in the past 14 days	4,019	0.226	0.233	0.220	-0.032	0.483
		(0.418)	(0.423)	(0.414)		
Was ill with a cough in the past 14 days	4,019	0.381	0.400	0.363	-0.077	0.087
		(0.486)	(0.490)	(0.481)		
Was ill with diarrhea in the past 14 days	4,019	0.170	0.178	0.163	-0.040	0.308
		(0.376)	(0.382)	(0.369)		
Duration of diarrhea in the past 14 days (if child had diarrhea)	683	3.407	3.487	3.323	-0.073	0.381
		(2.258)	(2.443)	(2.047)		
Received BCG vaccine	3,979	0.972	0.970	0.973	0.022	0.544
		(0.166)	(0.171)	(0.161)		
Received polio vaccine	3,972	0.973	0.973	0.972	-0.006	0.875
		(0.163)	(0.161)	(0.164)		
Received DPT-HB vaccine	3,968	0.915	0.918	0.912	-0.022	0.469
		(0.280)	(0.275)	(0.284)		
Received MMR or measles injection (9 months or older)	2,531	0.899	0.903	0.895	-0.027	0.484

	N	All	Control	Treatment	Normalized Difference	P-value
		(0.302)	(0.296)	(0.307)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample of children born in 2011 or later who are in the household. Standard deviations are in parentheses.

The primary females report that 82.1% of children are in good health, very good health, or excellent health (average health and poor health being the other two options). Despite this, 22.6% of children were ill with a fever, 38.1% were ill with a cough, and 17.0% of children were ill with diarrhoea during the past 14 days. Among children who had at least one episode of diarrhoea, the mean number of days with diarrhoea was 3.4. This suggests that illness is quite common among young children in the sample.

Vaccination rates are high—97.2% of children received a BCG vaccine, 97.3% received a polio vaccine, 91.5% received a DPT-HB vaccine, and 89.9% received an MMR or measles injection. While we do not have information on whether children completed the full vaccination course for DPT-HB or MMR—many of the children were too young to have completed either course—the data we do have suggest that most households in the sample are able to access critical early childhood vaccines for their children.

The largest normalized differences in Table 5.7 is -0.08, suggesting the two treatment arms are well balanced with respect to these characteristics. Of the nine p-values, none are significant at the 5% level. This again indicates that general child health and vaccination histories for children are similar in the treatment and control groups.

5.8 Antenatal and Postnatal Care

Take-up of antenatal care is extremely common: the primary female respondents report having attended at least one antenatal check-up for 99.8% of their children under the age of five. While a small share of women visited private hospitals or clinics (1.8%) or religious or voluntary hospitals (3.8%), a vast majority received antenatal care at a government health facility (94.3%). On average, women reported attending 4.1 antenatal visits, suggesting most make the recommended four antenatal visits (Kearns et al., 2014). The Focused Antenatal Care (FANC) model of antenatal care adopted by the MoHCDGEC in 2002 recommends that women attend their first antenatal visit prior to 16 weeks; the women who atteneded at least one antenatal check-up, on average, attend their first visit after 3.6 months (~15.7 weeks), and 53.6% attend their first visit before the 16 week recommendation for first visit timing.

Table 5.8: Maternal Health, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Received antenatal care	2,865	0.998	0.999	0.997	-0.049	0.258
		(0.042)	(0.027)	(0.052)		
Received antenatal care at home	2,860	0.001	0.000	0.002	0.064	0.081
		(0.032)	(0.000)	(0.045)		
Received antenatal care at a government health facility	2,860	0.943	0.927	0.959	0.141	0.061
		(0.231)	(0.261)	(0.197)		
Received antenatal care at a private hospital/clinic	2,860	0.018	0.023	0.013	-0.073	0.130
		(0.132)	(0.149)	(0.114)		
Received antenatal care at a religious or voluntary hospital	2,860	0.038	0.050	0.025	-0.131	0.089
		(0.191)	(0.219)	(0.158)		
Number of months pregnant at time of first ANC visit	2,860	3.637	3.677	3.598	-0.064	0.293
		(1.233)	(1.223)	(1.242)		
Number of ANC visits	2,860	4.117	4.036	4.196	0.121	0.019
		(1.326)	(1.261)	(1.381)		

	N	All	Control	Treatment	Normalized Difference	P-value
Blood pressure measured during antenatal care	2,843	0.821	0.802	0.839	0.096	0.096
		(0.384)	(0.399)	(0.368)		
Urine sample collected during antenatal care	2,851	0.698	0.672	0.724	0.112	0.048
		(0.459)	(0.470)	(0.447)		
Blood sample collected during antenatal care	2,855	0.958	0.952	0.965	0.066	0.148
		(0.200)	(0.215)	(0.184)		
Told about signs of pregnancy complications	2,839	0.617	0.613	0.622	0.019	0.721
		(0.486)	(0.487)	(0.485)		
Received tetanus injections	2,852	0.847	0.840	0.854	0.041	0.400
		(0.360)	(0.367)	(0.353)		
Received iron supplements	2,856	0.943	0.939	0.948	0.039	0.420
		(0.231)	(0.240)	(0.223)		
Had difficulty with vision during the daylight	2,865	0.055	0.059	0.051	-0.036	0.386
		(0.228)	(0.236)	(0.220)		
Suffered from night blindness	2,865	0.034	0.043	0.025	-0.103	0.015
		(0.181)	(0.204)	(0.155)		
Slept under a bednet	2,865	0.850	0.837	0.862	0.071	0.252
		(0.358)	(0.370)	(0.345)		
Took anti-malaria medication	2,854	0.895	0.884	0.906	0.071	0.167
		(0.306)	(0.320)	(0.292)		
Instructed to attend a follow up visit at nearest dispensary (after birth)	2,865	0.770	0.750	0.789	0.092	0.098
		(0.421)	(0.433)	(0.408)		
Place of delivery: Home	2,843	0.043	0.049	0.037	-0.058	0.264
		(0.203)	(0.216)	(0.190)		
Place of delivery: Government health facility	2,843	0.898	0.887	0.909	0.075	0.197
		(0.303)	(0.317)	(0.287)		
Place of delivery: Private hospital/clinic	2,843	0.021	0.019	0.022	0.020	0.663
		(0.143)	(0.138)	(0.147)		
Place of delivery: Religious or voluntary hospital	2,843	0.038	0.045	0.031	-0.072	0.268
		(0.191)	(0.207)	(0.174)		
Had a postnatal follow up within the first week of the child's birth	2,865	0.866	0.868	0.865	-0.009	0.849
		(0.340)	(0.339)	(0.342)		
Number of postnatal follow ups within 28 days of the child's birth	2,848	1.681	1.684	1.679	-0.004	0.937
		(1.395)	(1.379)	(1.411)		
Number of postnatal follow ups within 42 days of the child's birth	2,840	2.549	2.523	2.574	0.029	0.510
		(1.748)	(1.620)	(1.863)		
Visited at home by community health worker during the first few days after child	2,865	0.164	0.146	0.181	0.094	0.122
		(0.370)	(0.354)	(0.385)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample of all births in 2013 or later. Standard deviations are in parentheses.

For women who attended at least one antenatal visit, the 82.1% had their blood pressure monitored, 69.8% had a urine sample collected, 95.8% had a blood sample collected, 61.7% were informed about the signs of pregnancy complications, 84.7% received tetanus injections, and 94.3% received iron supplements. The high rate of iron supplementation is particularly heartening, given existing research linking iron supplementation during pregnancy to reduced maternal anaemia and iron deficiency in pregnancy (Pena-Rosas et al., 2015). Vision difficulties during daytime (5.5%) and night blindness (3.4%) were uncommon for the women in the sample.

Results indicated that 85.0% of women slept under a bed net while pregnant and 89.5% took anti-malarial medication during pregnancy, both of which were critically important, given that malaria is one of the leading causes of mortality in Tanzania (WHO, 2004) and it can additionally have potentially severe adverse consequences for the health of the foetus (WHO, 2014).

Government health facilities are by far the most common location for delivery to occur: 89.8% of births for children under five occurred at a government health facility as compared to just 4.3% at home, 2.1% at a private hospital or clinic, and 3.8% at a religious hospital. Women report that they attended a follow-up visits within a week of the birth for 86.6% of past births and, on average, had attended 1.7 postnatal visits within 28 days of the birth, and 2.5 postnatal visits within 42 days of the birth, figures which fall short of the three postnatal visits within the first 42 days recommended by guidelines from the Tanzanian Ministry of Health and Social Welfare (Ministry of Health & Social Welfare, 2010).

Though women in the sample report fairly high levels of adherence with the recommended antenatal visit programme—with respect to the number of visits and the services received (tests conducted, information received, iron supplementation, bed net use, antimalarial medication)—compliance is not universal. This suggests that there may be scope for an intervention to remind and encourage currently non-compliant households to improve their antenatal behaviours. Similarly, postnatal visits to a healthcare provider occur less frequently than recommended. To the extent that this non-compliance is driven by a lack of information, the mNutrition programme could help to improve postnatal visit attendance through the provision of information.

Antenatal and postnatal behaviour are similar in the two treatment groups. Of the 26 characteristics shown in Table 5.8, the largest normalized difference is 0.14, for whether women received antenatal care at a government health facility prior to past births. We are able to reject the null hypothesis of no difference between the treatment group mean and the control group mean for just three variables at the 5% level: the number of antenatal visits, whether a urine sample was collected during antenatal care, and whether the mother ever suffered from night blindness while pregnant. Further, these differences in means are small relative to the typical variation for each variable we observe in the sample; the largest normalized difference for these three characteristics is 0.12—for the number of antenatal visits—still less than half of the 0.25 threshold that would identify a problematic lack of balance.

5.9 HIV/AIDS Awareness

According to UNAIDS, the Iringa region had one of the highest HIV prevalence rates in Tanzania at the time of the 2011-2012 Demographic and Health Survey (UNAIDS, 2014). Though a number of the higher HIV prevalence districts that were a part of Iringa in 2011 have since become part of different administrative regions, estimated HIV prevalence among adults aged 15-49 in the three rural districts included in the quantitative study remains high: 7.6% in Iringa rural, 6.4% in Kilolo, and 13.4% in Mufindi. All three prevalence rates are substantially higher than the Tanzania wide estimate of 4.7% (USAID, 2016). Given the observational evidence indicating there is likely a negative relationship between maternal HIV status and child nutrition (Lartey et al., 2014), understanding perceptions and HIV related experiences for the women in the sample is critical. In addition, as a part of the general health information included in HPHB, the programme includes some content focused on the prevention of mother to child transmission of HIV (PMTCT) and encouraging testing for HIV. For both reasons, we therefore felt it was important to measure households' awareness of HIV/AIDS, HIV testing, and HIV treatment.

Module D, Part 2 of the household survey collects information from the primary female respondents on their awareness of HIV/AIDS. While we avoid asking directly about whether respondents, their partners, or anyone else in their household is HIV positive, the data are still useful for understanding attitudes and access to diagnostic and treatment options in the study context. Table 5.9 presents the information collected through this survey module.

Table 5.9: HIV/AIDS Awareness, by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Heard of illness called HIV/AIDS	2,833	0.991	0.992	0.990	-0.021	0.564
		(0.094)	(0.088)	(0.099)		
Tested for the HIV virus	2,833	0.986	0.984	0.987	0.026	0.572
		(0.118)	(0.124)	(0.112)		
Received results for the HIV virus test (of those tested)	2,793	0.993	0.990	0.996	0.069	0.094
		(0.084)	(0.100)	(0.065)		
Partner tested for the HIV virus	2,125	0.944	0.940	0.949	0.037	0.476
		(0.229)	(0.237)	(0.221)		
Heard of antiretroviral therapy	2,833	0.809	0.808	0.810	0.006	0.906
		(0.393)	(0.394)	(0.392)		
Knows where to receive antiretroviral therapy	2,292	0.728	0.712	0.744	0.073	0.147
		(0.445)	(0.453)	(0.437)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

HIV/AIDS awareness is extremely high among the primary females in the sample: 99.1% of women report having heard of HIV/AIDS and nearly as many (98.6%) report having been tested for the HIV virus. This is substantially higher than the 74.4% of women who reported ever having been tested for HIV during the 2011-2012 DHS AIDS indicators survey (AIS) in Iringa, but similar to the 95.2% of women who had given birth during the last two years that reported being tested during an antenatal visit in the same AIS. Of those tested, 99.3% received the results of their test. Women report that their partners are less likely to have been tested (94.4%), though this difference could also be driven by measurement error in the primary females' responses—women may not have perfect information about their partners' testing behaviour. The respondents are less likely to have heard of antiretroviral therapy (ARTs)—80.9% indicate they have heard of ARTs—and of those who heard of ARTs, just 72.8% would know where to access ARTs.

The sample is balanced with respect to all of the HIV/AIDS awareness indicators. The largest normalized difference is just 0.073 and none of the tests for differences between the treatment and control group means are significant at the 5% level.

5.10 Marriage and Fertility

A growing literature identifies how fertility and fertility preferences can affect child nutrition (Barcellos et al., 2014; Jayachandran and Kuziemko, 2011; Palloni, 2017): children born of their parents less desired sex receive fewer time and material resources, are more likely to have younger siblings and shorter subsequent birth intervals, and have lower BMI-for-age z-scores and increased mortality rates. To better understand fertility preferences in the study sample, the household survey includes a detailed set of questions on the marital history for primary females as well as the fertility history and desired subsequent fertility for both primary females and primary males. Table 5.10 presents the results for primary females; Table 5.11 does so for primary males.

5.10.1 Primary Females

Results indicate that 29.9% of the primary females are currently pregnant; those that are currently pregnant are, on average, 25.7 weeks into the pregnancy and have known about their pregnancy for 16.3 weeks. This suggests that the pregnant women in the sample found out about being pregnant at just 9.4 weeks' gestation.

Considering additional children, 60.2% of the women would like to have another child, with nearly all of the remaining women preferring not to have another child (34.3%). Among those who would like to have another child, the respondents would like to wait 4.5 years, on average, before conceiving their next child.

39.5% of the women who are not currently pregnant are currently using a contraceptive method; 79.9% of pregnant women plan on using a contraceptive method after their pregnancy. Injectable contraceptives are the most common method of contraception currently being used by the female respondents: 59.2% of those on contraception are using an injectable, 13.1% use an IUD, 9.0% are on the pill, only 3.4% report using male condoms and just another 0.3% are using female condoms, 0.6% have been sterilized

Women report being approximately midway through their desired fertility: they have 2.5 children and desire 2.4 additional children, on average. Among the desired future children, women have a slight preference for male children: they desire 0.8 male children, 0.7 female children, and 0.8 children for whom sex would not matter. Of the primary female respondents with a partner, 87.5% believe that their partner desires the same number of additional children.

Table 5.10: Desired fertility (female), by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Primary female is currently pregnant	2,833	0.299	0.302	0.296	-0.014	0.585
		(0.458)	(0.459)	(0.456)		
Gestation of pregnancy (weeks)	844	25.687	25.719	25.655	-0.008	0.902
		(7.796)	(7.987)	(7.607)		
Time since respondent found out about their pregnancy (weeks)	845	16.301	16.582	16.019	-0.064	0.394
		(8.778)	(8.878)	(8.678)		
Time respondent would like to wait before conceiving their next child (years)	1,780	4.457	4.477	4.438	-0.024	0.680
		(1.663)	(1.621)	(1.704)		
Would like to have another child	2,833	0.602	0.606	0.597	-0.019	0.661
		(0.490)	(0.489)	(0.491)		
Would not like to have another child	2,833	0.343	0.334	0.352	0.039	0.358
		(0.475)	(0.472)	(0.478)		
Is undecided about having another child	2,833	0.055	0.060	0.050	-0.041	0.324
		(0.228)	(0.237)	(0.219)		
Will use a contraceptive method after their pregnancy	782	0.799	0.784	0.815	0.078	0.259
		(0.401)	(0.412)	(0.389)		
Is using a contraceptive method	1,987	0.395	0.389	0.401	0.023	0.679
		(0.489)	(0.488)	(0.490)		
Method of contraceptive being used: IUD	785	0.131	0.131	0.132	0.002	0.984
		(0.338)	(0.338)	(0.338)		
Method of contraceptive being used: Injectables	785	0.592	0.529	0.653	0.253	0.002
		(0.492)	(0.500)	(0.477)		
Method of contraceptive being used: Pill	785	0.090	0.118	0.065	-0.186	0.027
		(0.287)	(0.323)	(0.246)		
Method of contraceptive being used: Condom	785	0.034	0.031	0.037	0.032	0.663
		(0.182)	(0.175)	(0.190)		
Method of contraceptive being used: Female Condom	785	0.003	0.005	0.000	-0.102	0.156
		(0.050)	(0.072)	(0.000)		
Method of contraceptive being used: Sterilization	785	0.006	0.008	0.005	-0.036	0.606
		(0.080)	(0.088)	(0.070)		
Number of living children	2,833	2.466	2.422	2.509	0.048	0.303
		(1.814)	(1.754)	(1.871)		

	N	All	Control	Treatment	Normalized Difference	P-value
Number of additional children wanted	1,860	2.357	2.348	2.366	0.015	0.814
		(1.255)	(1.274)	(1.237)		
Number of additional male children wanted	1,860	0.824	0.846	0.801	-0.049	0.342
		(0.923)	(0.906)	(0.939)		
Number of additional female children wanted	1,860	0.745	0.757	0.733	-0.027	0.597
		(0.882)	(0.880)	(0.883)		
Number of additional children wanted where the sex would not matter	1,860	0.788	0.744	0.832	0.072	0.138
		(1.219)	(1.225)	(1.212)		
Partner would like to have the same number of children	2,173	0.488	0.481	0.496	0.031	0.546
		(0.500)	(0.500)	(0.500)		
Partner would like to have more children	2,173	0.051	0.053	0.049	-0.016	0.745
		(0.220)	(0.224)	(0.217)		
Partner would like to have fewer children	2,173	0.019	0.019	0.020	0.011	0.802
		(0.138)	(0.135)	(0.140)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

The largest normalized difference in Table 5.10 is 0.25, the largest identified up to this point, for the likelihood that the woman—conditional on using a contraceptive method—is using an injectable. This suggests that women in treatment villages are substantially more likely to be using an injectable than those in control villages. However, the next largest normalized difference—for the likelihood that the respondent is using the pill—is of the opposite sign (-0.19). These two characteristics are also the only two where there is a significant difference in the mean value between treatment households and control households, which is consistent with the fact that there is no difference in the likelihood of currently being on any contraception across the two treatment groups.

5.10.2 Primary Males

The primary males in the sample seem to have accurate information about their partners' current pregnancy: 30.7% report having a wife/partner who is currently pregnant, the mean gestation for those with a pregnant partner is 25.2 weeks, and the males indicate that they found out about the pregnancy, on average, 15.2 weeks ago. The men in the sample, on average, find out about pregnancies just over half a week later than the primary females.

Regarding additional children, 61.0% of the primary males would like to have another child, a proportion that is nearly identical to the percent of primary females who desire another child. The primary males have, on average, 3.3 existing children and desire 2.4 additional children. The larger number of existing children for the primary males relative to the primary females in the sample is likely driven by a combination of polygamy and extramarital sex.

Men report a clear preference for future children to be male: of 2.4 additional children, they desire 0.9 male children, 0.7 female children, and 0.8 children where the sex would not matter. This desired male ratio of 1.4 (the desired number of male children over the desired number of female children) is over 25% higher than the desired male ratio reported by the primary female respondents, suggesting male preference is substantially stronger among adult males than adult females.

The same pattern of comparatively higher injectable use and lower use of birth control pills is reported by the primary male sample. However, males indicate much higher use of condoms: 64.2% of primary males report using male condoms compared to just 3.4% of primary females.

Table 5.11: Desired fertility (male), by mNutrition beneficiary status, Full Sample

	N	All	Control	Treatment	Normalized Difference	P-value
Wife/partner is currently pregnant	1,207	0.307	0.299	0.314	0.033	0.526
		(0.461)	(0.458)	(0.465)		
Gestation of pregnancy (weeks)	358	25.212	24.829	25.579	0.094	0.402
		(8.005)	(8.435)	(7.575)		
Time since respondent found out about their pregnancy (weeks)	354	15.229	14.929	15.503	0.067	0.519
		(8.514)	(8.560)	(8.487)		
Time respondent would like to wait before conceiving their next child (years)	785	3.925	3.998	3.844	-0.096	0.176
		(1.603)	(1.689)	(1.502)		
Would like to have another child	1,207	0.610	0.636	0.583	-0.108	0.114
		(0.488)	(0.481)	(0.493)		
Would not like to have another child	1,207	0.337	0.309	0.365	0.119	0.089
		(0.473)	(0.462)	(0.482)		
Is undecided about having another child	1,207	0.053	0.055	0.051	-0.016	0.785
		(0.224)	(0.228)	(0.221)		
Will use a contraceptive method after their pregnancy	348	0.727	0.692	0.761	0.156	0.164
		(0.446)	(0.463)	(0.427)		
Is using a contraceptive method	837	0.270	0.273	0.267	-0.011	0.876
		(0.444)	(0.446)	(0.443)		
Method of contraceptive being used: IUD	226	0.035	0.026	0.045	0.102	0.431
		(0.185)	(0.160)	(0.208)		
Method of contraceptive being used: Injectables	226	0.212	0.157	0.270	0.279	0.042
		(0.410)	(0.365)	(0.446)		
Method of contraceptive being used: Pill	226	0.062	0.087	0.036	-0.212	0.191
		(0.242)	(0.283)	(0.187)		
Method of contraceptive being used: Condom	226	0.642	0.670	0.613	-0.118	0.408
		(0.481)	(0.472)	(0.489)		
Method of contraceptive being used: Female Condom	226	0.009	0.000	0.018	0.191	0.153
		(0.094)	(0.000)	(0.134)		
Method of contraceptive being used: Sterilization	226	0.000	0.000	0.000		
		(0.000)	(0.000)	(0.000)		
Number of living children	1,207	3.282	3.145	3.418	0.102	0.167
		(2.677)	(2.366)	(2.949)		
Number of additional children wanted	800	2.418	2.418	2.417	-0.001	0.990
		(1.460)	(1.405)	(1.519)		
Number of additional male children wanted	800	0.921	0.875	0.971	0.093	0.204
		(1.037)	(1.024)	(1.051)		
Number of additional female children wanted	800	0.665	0.712	0.615	-0.113	0.133
		(0.862)	(0.925)	(0.787)		
Number of additional children wanted where the sex would not matter	800	0.831	0.832	0.831	-0.001	0.993
		(1.404)	(1.284)	(1.524)		
Partner would like to have the same number of children	1,207	0.486	0.475	0.496	0.042	0.474
		(0.500)	(0.500)	(0.500)		
Partner would like to have more children	1,207	0.031	0.033	0.028	-0.030	0.641
		(0.172)	(0.179)	(0.165)		
Partner would like to have fewer children	1,207	0.017	0.018	0.015	-0.027	0.673
		(0.128)	(0.134)	(0.121)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

As in Table 5.10, injectable use and the use of birth control pills are the two least balanced characteristics across the treatment and control groups in Table 5.11: the normalized differences for injectable use and use of birth control pills are 0.28 and -0.21, respectively; once again, they are of the opposite sign and there is no difference in the likelihood of currently using a contraceptive method or planning to use a contraceptive method after the current pregnancy. The remaining characteristics all have normalized differences well below the 0.25 threshold and there are no significant differences in means between the treatment group and the control group.

6 Baseline Data: The Household Level Randomization

To answer research question 5 we conducted a household level randomization among treatment households where the primary male and primary female respondents were both interviewed and they each owned their own mobile phone. The random assignment of eligible households to the primary female only treatment group (T-F)—where mNutrition content is just sent to the mobile phone of the primary female respondent—or to the primary male and primary female treatment group (T-F+M)—where mNutrition content is sent to the mobile phones of both the primary female and the primary male—ensures that we can estimate an unbiased answer to research question 5.

However, we can only answer research question 5 for the sub-sample of households that were eligible for this randomization. If households where the primary female and primary male both own different mobile phones are observably or unobservably different from those that do not, the treatment effects we estimate may not generalize to households that do not meet the eligibility criteria for the household level randomization. In this section, we therefore begin by comparing observable characteristics between the full sample of households and the sub-sample that were deemed eligible to participate in the household level randomization. It is worth emphasizing that differences in observable characteristics between the two samples would not suggest that our estimates of the effect of sending the mNutrition content to the primary female and primary male, as compared to just the phone of the primary female, are biased. Instead, we would simply need to be cautious when extrapolating the effects we estimate to broader segments of the population in Tanzania. To test for differences in an observable characteristic between the two samples we run an Ordinary Least Squares regressions of the characteristic on an indicator for whether the household was eligible for the household randomization, clustering standard errors at the village level. The p-value from the t-test of no difference between the two samples is the measure we use to assess balance.

In addition to testing for differences in observable characteristics between the full sample and the sub-sample eligible for the household level randomization, we also explore whether—for eligible households—the household randomization was successful. To do so, we follow the same procedures as in Section 5, calculating both the p-value from a test of the null hypothesis of no difference in means between the T-F and T-F+M groups, and we compute the normalized difference. Small normalized differences and few p-values below 0.05 will indicate that the household level randomization was successful at identifying two observably similar groups.

Because the sample size for the household-level randomization is substantially smaller, for some characteristics there are fewer than 20 individuals with non-missing values. Instead of showing the true observation count for these variables, we display a “<20” symbol in the observation count (N) column of the table.

The tables and full text describing the comparison between the main sample and the sample eligible for the household level randomization can be found in Annex F. In this section we focus on briefly discussing key differences and similarities between the two samples. Sub-group means and balance tests between the T-F and T-F+M groups—those eligible for the household level randomization—are also in Annex F, with the text in this section again limited to a summary.

6.1 Household Level Randomization: Comparison with the Overall Sample and Balance Across Sub-Treatment Groups

6.1.1 Comparing Households Eligible for the Household Randomization to the Overall Sample

Table 7.1, Table 7.2, and Table 7.3 show means and standard deviations for household demographic characteristics, primary outcomes, and secondary outcomes. In each table, the first column presents the

number of observations with a non-missing value in the full sample, the second column presents the mean and standard deviation (shown below the mean in parentheses) for the characteristic in the full sample, the third column displays the number of households that were eligible for the household level randomization and had a non-missing value, the fourth column presents the mean and standard deviation for the characteristic in the sample that was eligible for the household randomization, and the fifth column shows the p-value from a t-test of the null hypothesis of no difference between the means for the two samples.

Of the total number of treatment households, 276 households were identified as being eligible for the household level randomization, just under 20% of the total number. Table 7.1 indicates that eligible households are observably different from the full sample. In terms of demographic characteristics, they are significantly less likely to be female headed, more likely to have a head of household with some formal education, more likely to have a head of household and a primary female that are married and monogamous, and they have progress out of poverty scores that are significantly higher on average. Only one of the differences is in the primary and secondary outcomes (male nutrition knowledge) is statistically significantly different from zero; therefore, despite being observably wealthier than the overall sample, the sub-sample eligible for the household randomization is similar to the full sample with respect to child nutrition outcomes, IYCF practices, and nutrition knowledge. Still, the differences in demographic characteristics potentially question the external validity of our estimates to research question 5. Ex-ante, it is not clear how these differences should affect the return to participation in the mNutrition programme. On the one hand, wealthier, better educated households may find it easier to understand the information in the SMS messages or they may be more likely to have the resources necessary to act on some of the messaging; conversely, it is possible that wealthier households may already have the information contained in the mNutrition programme, and their nutrition related behaviours may be more likely to be compliant with the actions suggested by the programme.

6.1.2 Household Randomization: Testing for Balance Across Sub-Treatment Arms

The remainder of Section 6 tests for balance in observable characteristics across the two sub-treatment arms (T-F and T-F+M) within the sub-sample of households that were eligible for the household randomization. As mentioned above, we assess balance both by calculating the normalized difference and by testing for a difference in means for each characteristic. To avoid overburdening readers, we limit the discussion in the text to household demographic characteristics, primary outcomes and child anthropometry, and secondary outcomes, and place the all of the tables displaying the balance tests for the household randomization in Annex F.

By design, half of the 276 eligible households were allocated to the female only sub-treatment arm and the other half were assigned to the male and female sub-treatment arm. Of the 28 demographic characteristics tested, we are able to reject the null hypothesis of no difference between the means for the sub-treatment groups just once, and there is no clear pattern in the observed differences between the two groups. Similarly, there is good balance across the two groups with respect to the primary outcomes and child anthropometry: none of the differences are significant at the 5% level and only one normalized difference is greater than 0.25, children in the male and female sub-treatment arm more likely to be stunted than those in the female only arm.

For some of the IYCF indicators that are missing for all children outside of a narrow age range, the sample size drops below 50 total observations across both groups, making it difficult to interpret the difference in means tests with much confidence. That said, the normalized differences, which are less sensitive to changes in sample size, do not suggest that there is important imbalance in any of the IYCF practices. The IYCF knowledge and beliefs measures, shown in Table 7.6, are also balanced across the sub-treatment groups.

7 Conclusion

The baseline data collection generated a great deal of relevant data for the quantitative evaluation, which we summarize in detail in Sections 5 and 6. Here we highlight some of the most important findings and discuss the implications that the sample characteristics and the baseline balance in observables are likely to have for the quantitative evaluation.

7.1 Sample Characteristics

The 2,833 households surveyed during the baseline household survey are economically disadvantaged. For example, Table 5.1 indicates that based on the progress out of poverty index (PPI), 53.1% of sample households are below 150% of the national poverty line and 68.3% are living on less than $2.00 per day in 2005 US dollars.

The sample households have 5.3 members on average, with a mean difference in age between the primary female and primary male respondent of 11.6 years—primary females are 27.0 years of age and primary males are 38.6 years of age, on average. Overwhelmingly, the respondents list crop production as their principal activity: 73.1% of primary males and 76.8% of primary females report crop production as their main activity. Formal school attendance is high: among children between the ages of 5 and 20, 66.2% are attending school.

75.7% of the sample have access to improved sanitation and 82.4% have access to improved sources of drinking water. Given the increasing body of evidence on the critical role played by water, sanitation, and hygiene (WASH) in determining child mortality, morbidity, and growth (Pruss-Ustun et al., 2008; Spears, 2013), ensuring universal access to safe water and improved sanitation offers one potential pathway to increase nutrition outcomes for the children in the sample.

Child height, widely considered to be the best early childhood measure of overall health and nutrition (Black et al., 2013; Currie and Vogl, 2013), is significantly below the age and sex-specific growth standards as defined by the WHO (WHO, 2006). On average, the sample children measured have a height-for-age z-score of -1.4, suggesting they are nearly a standard deviation and a half below the reference group mean with nearly a third of the children (29.6%) being stunted (HAZ below -2). This both highlights the likelihood of substandard nutrition behaviour among the sample households and underscores the importance of identifying ways to improve early childhood nutrition for the children in the sample.

Direct reports on existing Infant and Young Child Feeding (IYCF) practices and antenatal and postnatal care for the sample offer some suggestion for what nutrition and health related behaviours might be the most useful for the mNutrition informational content to target. 76.4% of the women in the sample initiated breastfeeding within one hour of birth for their children born in the 24 months preceding the interview date but just 1.3% of children are exclusively breastfed for the first six months of their lives. With non-universal access to clean drinking water and improved sanitation, the low rate of exclusive breastfeeding is particularly problematic for child nutrition. Child dietary diversity for children 6-23 months of age is low: just 21.4% of the children in this age group consume from four or more food groups. Similarly, minimum meal frequency is satisfied by under half (45.5%) of the children between 6 and 23 months of age. Women's dietary diversity is also below desired levels, with a mean WDDS of 3.5 for the primary female respondents.

Vaccination rates are extremely high, albeit not universal, for children under five: 97.2% received a BCG vaccine, 97.3% received a polio vaccine, 91.5% were given a DPT-HB vaccine, and 89.9% were given an MMR or measles vaccine. Self-reported use of antenatal care for the primary females in the sample is also quite high. 99.8% of women reported receiving some antenatal care, with a mean number of antenatal visits of 4.1. On average, women attended their first antenatal visit when they were 3.6 months pregnant and 94.3% received iron supplements during their pregnancy. Anti-malarial medication was taken by 89.5%

of women during pregnancy and 85.0% of primary females reported sleeping under a bednet while pregnant.

Adherence to guidelines for postnatal care is less common among the primary females in the sample. Women report that they attended a follow-up visits within a week of the birth for 86.6% of past births and that they, on average, had attended 1.7 postnatal visits within 28 days of the birth, and 2.5 postnatal visits within 42 days of the birth. These figures fall short of the three postnatal visits within the first 42 days recommended by guidelines from the Tanzanian Ministry of Health and Social Welfare (Ministry of Health & Social Welfare, 2010).

On the whole, there is scope for improvement in both antenatal and postnatal behaviour for the households in the sample. Similarly, IYCF practices, especially exclusive breastfeeding, children's dietary diversity, and meal frequency fall well below recommended levels. The mNutrition content that addresses these key behaviours may therefore prove to be particularly beneficial for the children and households in our sample.

There is also substantial room for improvement with respect to IYCF knowledge and beliefs for the primary females and primary males in our sample. On average, primary female respondents answered just 55.9% of the 11 IYCF knowledge questions correctly. Primary male respondents fared even worse, answering only 48.5% of the questions correctly. Together with the reported gap in IYCF behaviours discussed in the previous paragraphs, the low levels of IYCF knowledge among households in the sample suggest that the provision of nutrition related information may be a particularly productive way to improve nutrition outcomes for the children in the sample.

Though there are important gaps in nutrition knowledge and behaviours for the sample households, the mNutrition is designed to provide information through mobile phones. For the intervention to be successful, it is therefore imperative that sample households have access to mobile phones. Fortunately, Table 5.5 and Table 5.6 suggest that access to and use of mobile phones is extremely prevalent in our sample. 58.9% of primary female respondents and 90.7% of primary male respondents own a mobile phone. Though making and receiving phone calls are the most common way that the respondents interact with mobile phones, 83.5% of primary males and 62.4% of primary females used a mobile phone to receive an SMS message during the fourteen days preceding the date of the interview. Most of the primary female respondents also report being willing to trust health and nutrition information that they receive through automated SMS messages; 82.0% of primary female indicated that they would trust health or nutrition information that they receive through an automated text message from an NGO, private organization, religious or voluntary organization, or the government.

7.2 Baseline Balance in Observable Characteristics

The quantitative sample appears to consist largely of the type of household we should expect to benefit from an intervention like the mNutrition programme: there are important gaps in nutrition knowledge, behaviour, and outcomes, but households have high levels of access to mobile phones and they report being likely to trust health and nutrition information that they receive through a mobile phone. To assess how likely it is that the quantitative evaluation will be able to accurately estimate the causal effect of access to the mNutrition programme on outcomes[48] we now turn to summarizing the observed balance in baseline characteristics between the treatment and control groups, as well as between the two sub-treatment groups designed to answer research question 5 (T-F and T-F+M).

We assess balance using two different measures: the normalized differences between the treatment and control group distributions and p-values from tests of the null hypotheses of no difference in means between the treatment and control groups. We follow Imbens (2015) and interpret normalized differences below 0.25 in absolute value as being indicative of sufficient balance for the variable being tested. To be consistent with our plans for the impact evaluation after collecting endline data, we also treat p-values

[48] Or registration for the programme among household compliers (LATE for compliers) as discussed in Section 3.5.2.

below 0.05 as evidence of imbalance in the characteristic being tested—though ultimately we prioritize the scale and sample size free normalized difference and we recognize that we should expect to observe roughly one out of every twenty tests have a p-value below 0.05 simply by chance.

7.2.1 Village Level Randomization

The baseline balance in observable characteristics between the treatment and control groups is shown in Table 5.1 through Table 5.11. Across all elven tables, the baseline balance is good. This is true regardless of whether we use the normalized difference or the number of p-values below 0.05 to assess balance.

The baseline report does not make any adjustments for multiple hypothesis testing—that is, adjusting for the fact that the likelihood of rejecting the null hypothesis at any level α when conducting multiple tests is typically increasing in the number of tests. Suggested corrections for this over rejection of the null due to multiple inference range from a Bonferroni correction that adjusts the required p-value for rejecting the null hypothesis by scaling the original significance level by the number of tests m ($\alpha_{Bonf} = \frac{\alpha}{m}$), reducing the number of tests by generating a summary index that combines the data from the individual indicators, or step-down methods for adjusting p-values to control for the familywise error rate (FWER)[49] using the actual data (Romano and Wolf, 2005; Kling et al., 2007; Anderson, 2008). While the endline impact analysis will use both the summary index and step-down p-value adjustment methods, we elect to make no formal adjustments for multiple hypothesis testing during the baseline analysis. Instead, we simply note that any adjustment for multiple hypothesis testing would reduce the number of tests with a p-value below our significance level of $\alpha = 0.05$.

Of the 204 characteristics tested, only two (or .9%) have normalized differences that are above the 0.25 threshold. These normalized differences are observed for the likelihood that children aged 12-15 months are still breastfed and for the likelihood that women report using injectables as contraception. In both cases, the other indicators from the broader topics—IYCF practices and contraceptive use—show no signs of imbalance and are not consistently signed in the same direction as the differences for breastfeeding of 12 to 15 month-old children and use of injectables.

We reject the null hypothesis of no difference in means between treatment households and control households for 12 of the 204 characteristics tested. This is a rejection rate of 0.06%, almost exactly what we should expect to find by chance; with a significance level of 0.05, we should expect to falsely reject the null hypothesis, given that it is true, 5% of the time.

Based on both the number of normalized differences above 0.25 and the number of tests of no difference in means between the treatment and control group with a p-value below 0.05, the village level randomization was highly successful at balancing observable characteristics across the treatment and control groups. This suggests that estimating the causal impact of the mNutrition programme on nutrition knowledge, behaviours, and outcomes will be possible through any of the empirical strategies discussed in Section 3.5.

7.2.2 Household Level Randomization

In the same way as we tested for balance between the treatment groups defined by the village level randomization, we can also test for balance between the sub-treatment groups generated by the household level randomization: the T-F group that was assigned to just receive the mNutrition content on the mobile of the primary female and the T-F+M group that was assigned to receive the mNutrition content on the mobile phones of both the primary female and the primary male respondents. The variable level balance tests are shown in Table 7.4 through Table 7.14 in132Annex F.

[49] The FWER is defined as the probability of rejecting at least one true null hypothesis.

As with the village level randomization, balance across the sub-treatment groups defined by the household level randomization is extremely good; just nine of the 204 variables tested have normalized differences above the 0.25 threshold in absolute value and nine variables have p-values from the test of no difference in means between the treatment and control group that are below 0.05. Therefore, we observe a significant difference at the 5% level 4.4% of the time—in line with what we should expect to observe by chance—and large normalized differences occur with the same likelihood.

After collecting the endline data, we will ensure that empirical estimates based on the household level randomization are robust to including controls for the characteristics that were not well balanced across the sub-treatment groups. Given the paucity of imbalance across the two sub-groups, we expect that these differences are driven principally by chance, and therefore that the treatment effects we estimate will not be importantly affected by the inclusion of controls for characteristics that displayed some baseline imbalance.

7.2.3 Limitations of the Study

While the rapid spread of access to and ownership of mobile phones across the developing world has generated a potentially low-cost opportunity for disseminating information to individuals and households, there remain practical challenges for mobile phones to be an effective means of improving nutrition knowledge, behaviours, or nutrition outcomes. In contrast with more typical in-person methods of conducting behaviour change communication whereby programme staff are physically available to deliver content to beneficiaries, to work, mobile phone-based information interventions need to ensure that targeted mobile phone numbers are still in use (working, charged, and accessible), and that the desired beneficiary has the time and desire to read the delivered content. While our data indicate that mobile phone ownership and accessibility are high among the study population, the baseline data cannot provide information on the frequency of mobile phone or subscriber identity module switching, or on the probability that delivered messages were read by the targeted recipients, both of which therefore represent potential barriers to the effectiveness of the mNutrition programme.

A second limitation of the study is the reduced statistical power to detect differences in outcomes for the sub-sample eligible for the household-level randomization. Despite not achieving the desired sample size for research questions 1-4, Table 3.1 suggests we should have better statistical power than anticipated for the HAZ outcome—because of a lower intracluster correlation and a smaller standard deviation in the baseline data—and that we will be able to detect differences in WDDS as small as 0.27 food groups and differences in a combined IYCF practices index as small as 0.06 standard deviations. Unfortunately, male response rates were sufficiently low that the sample eligible for the household-level randomization is limited: just 276 households in treatment villages were identified as being eligible. The implied minimum detectable effects for research question iv5 are therefore substantially larger than the corresponding effect sizes for the village-level regression: 0.484 for HAZ, 0.533 food groups for WDDS, and 0.267 standard deviations for the IYCF practices index.

7.2.4 Next Steps

The quantitative evaluation will continue through December of 2018, with preparations for the endline survey beginning in the summer of 2018 and endline data collection activities starting in October of 2018. Between now and the end of the evaluation period the mNutrition programme will operate as discussed in Section 2: registered individuals will receive SMS-based nutrition and health information tailored to their pregnancy gestation or their child's age. Given the nearly perfect level of initial take-up of the mNutrition programme among treatment households (99.6%), we expect there to be consistently high exposure to the content throughout the study period for the treatment group. When the endline data has been collected and cleaned, we will conduct the analysis to estimate the causal impacts of the mNutrition programme using the outcomes and methods detailed in Section 3.5.

While we encountered some surprises during the baseline activities (e.g. a lower than expected fertility rate; the scarcity of households where both the primary male and primary female each have their own mobile phone), the evaluation design has held up well throughout the field work, treatment offers, and the subsequent analysis of baseline data. We anticipate that the evaluation will provide well-powered estimates of the causal impacts of the mNutrition programme for the primary and secondary research questions, and we look forward to presenting and describing the results in the quantitative impact evaluation report in October of 2019.

References

Aker, J., Burrell, J., Ghosh, I. (2016) 'The promise (and pitfalls) of ICT for agriculture initiatives' in *Agricultural Economics* 47(S1).

Alderman, H. (2006) 'Long-term consequences of early childhood malnutrition' in *Oxford Economic Papers* 58(3), 450–474.

Anderson, ML. (2008) 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry preschool, and early training projects' in *Journal of the American Statistical Association* 103(484),1481–1495.

Barcellos, SH., Carvalho, LS., Lleras-Muney, A. (2014) 'Child gender and parental investments in India: Are boys and girls treated differently?' in *American Economic Journal: Applied Economics* 6(1),157–189.

Barker, DJ., Winter, PD., Osmond, C., Margetts, B., Simmonds, SJ. (1989) 'Weight in infancy and death from ischaemic heart disease' in *Lancet* 2(8663), 577–580.

Barnett, I, Scott, N, Batchelor, S, and Haddad, L. (2016) 'Dial `N' for Nutrition? A Landscape Analysis of What We Know About m-Nutrition, m-Agriculture and m-Development' *IDS Working Paper 481.*

Barnett, I., Srivastava, S., and J. Gordon. (2017) 'Mobile phones, nutrition and health in Tanzania: Initial exploratory qualitative study report.' External evaluation of mobile phone technology-based nutrition and agriculture advisory services in Africa and South Asia. Brighton: IDS.

Bertrand, M., Duflo, E., Mullainathan, S. (2004) 'How much should we trust differences-in-differences estimates?' in *The Quarterly Journal of Economics* 119(1), 249–275.

Bhutta, ZA., Ahmed, T., Black, RE., Cousens, S., Dewey, K., Giugliani, E., Haider, BA., Kirkwood, B., Morris, SS., Sachdev, HPS., Shekar, M. (2008) 'What works? Interventions for maternal and child undernutrition and survival' in *The Lancet* 371(9610): 417-440.

Bhutta, ZA., Das, JK., Rizvi, A., Gaffey, MF., Walker, N., Horton, S., Webb, P., Lartey, A., Black, RE. (2013) 'Evidence-based interventions for improvement of maternal and child nutrition: What can be done and at what cost?' in *Lancet* 382(9890), 452-477.

Black, RE., Allen, LH., Bhutta, ZA., Caulfield, LE., de Onis, M., Ezzati, M., Mathers, C., Rivera, J. (2008) 'Maternal and child undernutrition: Global and regional exposures and health consequences' in *The Lancet* 371(9608), 243–260.

Black, RE., Victoria, CG., Walker, SP., Bhutta, ZA., Christian, P., de Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R., Uauy, R., The Maternal and Child Nutrition Study Group (2013) 'Maternal and child undernutrition and overweight in low-income and middle-income countries' in *The Lancet* 382(9890), 427–451.

Boyle, EH., King, M., Sobek, M. IPUMS-Demographic and Health Surveys: Version 4.1 [dataset]. Minnesota Population Center and ICF International, 2017.

Bruhn, M., McKenzie, D. (2009) 'In pursuit of balance: Randomization in practice in development field experiments' in *American Economic Journal: Applied Economics* 1(4), 200–232.

Christian, P., Lee, SE., Donahue, AM., Adair, LS., Arifeen, SE., Ashorn, P., Barros, FC., Fall, CH., Fawzi, WW., Hao, W., Hu, G., Humphrey, JH., Huybregts, L., Joglekar, CV., Kariuki, SK., Kolsteren, P., Krishnaveni, GV., Liu, E., Martorell, R., Osrin, D., Persson, LA., Ramakrishnan, U., Richter, L., Roberfroid, D., Sania, A., Ter Kuile, FO., Tielsch, J., Victoria, CG., Yajnik, CS., Yan, H., Zeng, L., Black, RE. (2013) 'Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low- and middle-income countries' in *International Journal of Epidemiology* 42(5),1340–1355.

Cole, S., Fernando, A. (2016) '`Mobile'izing agricultural advice: Technology adoption, diffusion and sustainability' *HBS Working Paper.*

Concern Worldwide and the Centre for Research on the Epidemiology of Disasters (UCL). (2015). 'Bringing nutrition actions to scale in Iringa, Njombe and Mbeya regions of Tanzania: In-depth analysis of the factors associated with stunting.' *Study Report Version 3.*

Courtois, P., Subervie, J. (2015) 'Farmer bargaining power and market information services' in *American Journal of Agricultural Economics* 97(3), 953-977.

Currie, J., Vogl, T. (2013) 'Early-life health and adult circumstance in developing countries' in *Annual Review of Economics* 5, 1–36. doi: 10.1146/annurev-economics-081412-103704. URL http://www.annualreviews. org/doi/abs/10.1146/annurev-economics-081412-103704.

Demographic and Health Surveys (2016) [Datasets]. Calverton, Maryland: ICF International.

de Onis, M., Dewey, G., Borghi, E., Onyango, A., Blossner, M., Daelmans, B., Piwoz, E., Branca, F. (2013) 'The World Health Organization's global target for reducing childhood stunting by 2025: rationale and proposed actions' in *Maternal & Child Nutrition* 9(S2), 6-26.

Duflo, E., Glennerster, R., Kremer, M. (2007) 'Using randomization in development economics research: A toolkit' in Schultz, TP., Strauss, J. (eds.) *Handbook of Development Economics,* Vol. 4. Elsevier Science Ltd.: North Holland.

Fafchamps, M., Minten, B. (2012) 'Impact of SMS-based agricultural information on Indian Farmers' in *World Bank Economic Review* 26(3), 383-414.

Gilligan, DO., Hidrobo, M., Hoddinott, JF., Peterman, A., Roy, S., Schwab, B., Buller, A., Heise, L. (2015) 'Expanding lessons from a randomized impact evaluation of cash and food transfers in Ecuador and Uganda' in *3ie Grantee Final Report.* International Initiative for Impact Evaluation (3ie): New Delhi, India.

Fisher, RA. (1935) 'The Design of Experiments' First edition. London: Oliver and Boyd

Flax, V., Negerie, M., Ibrahim, AU, Leatherman, S., Daza, E., Bentley, M. (2014) 'Integrating group counseling, cell phone messaging, and participant-generated songs and dramas into a microcredit program increases Nigerian women's adherence to international breastfeeding recommendations' in *Journal of Nutrition* 144, 1120-1124.

Food and Agriculture Organization of the United Nations. (2014) *FAOSTAT Statistics Database.* Rome: FAO.

Food and Agriculture Organization of the United Nations (2016) *Minimum Dietary Diversity for Women: A Guide to Measurement*. Technical report, FAO and FHI 360.

Gechter, M. (2016) 'Generalizing the results from social experiments: Theory and evidence from Mexico and India' *Working Paper.*

Gelli, A., Margolies, A., Santacroce, M., Sproule, K., Theis, S., Roschnik, N., Twalibu, A., Chidalengwa, G., Cooper, A., Moorhead, T., Gladstone, M., Kariger, P., Kutundu, M. (2017) 'Improving child nutrition and development through community-based childcare centres in Malawi- The NEEP-IE study: study protocol for a randomized controlled trial' in *Trials* 18(1), 284.

Gluckman, PD., Hanson, MA. (2004) 'Living with the past: Evolution, development, and patterns of disease' in *Science* 305(5691), 1733–1736.

Greevy, R., Silber JH., Cnaan, A., Rosenbaum, PR. (2004) 'Randomization inference with imperfect compliance in the ACE-inhibitor after Anthracycline randomized trial' *Journal of the American Statistical Association* 99(465), 7-15.

Guiteras, RP., Levine, DI., Polley, TH. (2016) 'The pursuit of balance in sequential randomized trials' in *Development Engineering* 1(1), 12–25.

Hildebrandt, N., Nyarko, Y., Romagnoli, G., Soldani, E. (2015) 'Price information, inter-village networks, and `bargaining spillovers': Experimental evidence from Ghana' *Working Paper.*

Hoddinott, J., Behrman, JR., Maluccio, JA., Melgar, P., Quisumbing, AR., Ramirez-Zea, M., Martorell, R. (2013) 'Adult consequences of growth failure in early childhood' in *The American Journal of Clinical Nutrition* 98(5), 1170–1178.

Holland, PW. (1986) 'Statistics and causal inference' in *Journal of the American Statistical Association* 81(396), 945–960.

Huybregts, L., Becquey, E., Zongrone, A., Le Port, A., Khassanova, R., Coulibaly, L., Leroy, JL., Rawat, R., Ruel, MT. (2017) 'The impact of integrated prevention and treatment on child malnoutrition and health: the PROMIS project, a randomized control trial in Burkina Faso and Mali' in *BMC Public Health* 17(237).

Imbens, GW. (2015) 'Matching methods in practice' in *Journal of Human Resources* 50(2), 373–419.

Imbens, GW., Rosenbaum, PR. (2005) 'Robust, accurate confidence intervals with a weak instrument: quarter of birth and education' in *Journal of the Royal Statistical Society, Series A* 168(Part 1), 109-126.

Imbens, GW., Rubin, DB. (2015) 'Causal Inference for Statistics, Social, and Biomedical Sciences' Cambridge University Press.

International Labour Organization. (2017) *ILOSTAT database*. Data retrieved in November 2017.

Jayachandran, S., Kuziemko, I. (2011) 'Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India' in *The Quarterly Journal of Economics* 126(3), 1485–1538. URL http://ideas. repec.org/a/oup/qjecon/v126y2011i3p1485-1538.html.

Jiang, H., Li, M., Wen, LM, Hu, Q., Yang, D., He, G., Baur, L., Qian, X. (2013) 'Effect off a short message service intervention on infant feeding practices in Shanghai, China: a prospective, community-based, controlled study' in *The Lancet* 382(P4).

Kearns, A., Hurst, T., Caglia, J., Langer, A. (2014) *Focused Antenatal Care in Tanzania: Delivering Individualized, Targeted, High-Quality Care*. Technical report, Maternal Health Task Force.

Kennedy, G., Ballard, T., Dop, MC. (2013) *Guidelines for Measuring Household and Individual Dietary Diversity*. Technical report, Food and Agriculture Organization of the United Nations.

Kling, JR., Liebman, JB., Katz, LF. (2007) 'Experimental analysis of neighborhood effects' in *Econometrica* 75(1), 83–119. URL http://ideas.repec.org/a/ecm/emetrp/v75y2007i1p83-119.html.

Labrique, A. Vasudevan, L., Kochi M. (2013) '12 common applications and a visual framework' in *Global Health Science* 1(2), 160-171.

Lartey, A., Marquis, GS., Mazur, R., Perez-Escarnilla, R., Brakohiapa, L., Ampofo, W., Sellen, D., Adu-Afarwuah, S. (2014) 'Maternal HIV is associated with reduced growth in the first year of life among infants in the eastern region of Ghana: The research to improve infant nutrition and growth (RIING) project' in *Maternal and Child Nutrition* 10(4), 604–616.

Leroy, JL., Ruel, M., Habicht, JP., Frongillo, EA. (2014) 'Linear growth deficit continues to accumulate beyond the first 1000 days in low- and middle-income countries: Global evidence from 51 national surveys' in *The Journal of Nutrition* 144, 1460-1466.

McKenzie, D. (2012) 'Beyond baseline and follow-up: The case for more t in experiments' in *Journal of Development Economics* 99(2), 210–221.

Ministry of Health and Social Welfare (2010) *National Postpapost Care Guidelines*. Technical report, Reproductive and Child Health Section of the Ministry of Health and Social Welfare.

Nglazi, M., Bekker, L., Wood, R., Hussey, G., Wiysonge, C. (2013) 'Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review' *in BMC Infectious Diseases* 13(566).

Olney, DK., Bliznashka, L., Pedehombga, A., Dillon, A., Ruel, MT., Heckert, J. (2016) 'A 2-year integrated agriculture and nutrition program targeted to mothers of young children in Burkina Faso reduces underweight among mothers and increases their empowerment: A cluster-randomized controlled trial' *in The Journal of Nutrition* 146(5), 1109-1117.

Palloni, GE. (2017) 'Childhood health and the wantedness of male and female children' in *Journal of Development Economics* 126(C), 19–32.

Pelletier, DL., Frongillo Jr, EA., Schroeder, DG., Habicht, JP. (1995) 'The effects of malnutrition on child mortality in developing countries' in *Bulletin of the World Health Organization* 73(4), 443–448.

Peña-Rosas, JP., De-Regil, LM., Garcia-Casal, MN., Dowswell, T. (2015) 'Daily oral iron supplementation during pregnancy' in *Cochrane Database Systematic Reviews* 22(7).

Pew Research Center. (2015) 'Cell phones in Africa: Communication lifeline' *Technical Report.*

Population and Housing Census. (2012) *Iringa Region: Basic Demographic and Socio-Economic Profile* Technical report, National Bureau of Statistics of the Government of Tanzania.

Pruss-Ustun, A., Bos, R., Gore, F., Bartram, J. (2008) *Safer Water, Better Health: Costs, Benefits, and Sustainability of Interventions to Protect and Promote Health*. Technical report, World Health Organization.

Rokicki, S., Cohen, J., Salomon, J., Fink, G. (2017) 'Impact of a text-messaging program on adolescent reproductive health: A cluster randomized trial in Ghana' in *American Journal of Public Health* 107(2), 298-305.

Romano, JP., Wolf, M. (2005) 'Stepwise multiple testing as formalized data snooping' in *Econometrica* 73(4), 1237–1282.

Rosenbaum, P. (2002) 'Covariance adjustment in randomized experiments and observational studies' in *Statistical Science* 17(3), 286-327.

Ruel, MT. (2010) 'The Oriente Study: Program and policy impacts' in *Journal of Nutrition* 140(2), 415–418.

Ruel, MT., Menon, P., Habicht, J-P., Loechl, C., Bergeron, G., Pelto, G., Arimond, M., Maluccio, J., Michaud, L., Hankebo, B. (2008) 'Age-based preventive targeting of food assistance and behaviour change and communication for reduction of childhood undernutrition in Haiti: A cluster randomised trial' in *The Lancet* 371(9612), 588–595.

Schroeder, DG., Martorell, R., Rivera, JA., Ruel, MT., Habicht, JP. (1995) 'Age differences in the impact of nutritional supplementation on growth' in *The Journal of Nutrition* 125(4), 1051S-1059S.

Small, DS., Ten Have TR., Rosenbaum, PR. (2008) 'Randomization inference in a group-randomized trial of treatments for depression: Covariate adjustment, noncompliance, and quantile effects' in *Journal of the American Statistical Association* 103(481), 271-279.

Spears, D. (2013) *How Much International Variation in Child Height Can Sanitation Explain*. Technical report, Policy Research Working Paper.

Svensson, J., Yanagizawa, D. (2009) 'Getting prices right: the impact of the market information service in Uganda' in *Journal of the European Economic Association* 7(2-3), 435-445.

UNAIDS. (2014) *Developing Subnational Estimates of HIV Prevalence and the Number of People Living with HIV: Tanzania*. Technical report, UNAIDS.

UNICEF. (2016). 'UNICEF Annual Report 2016: The United Republic of Tanzania.' New York, NY: UNICEF.

USAID. (2016) *HIVAIDS in Tanzania*. Technical report, USAID.

Victoria, CG., Adair, L., Fall, C., Hallal, P., Martorell, R., Richter, L., Sachdev, HS. (2008) 'Maternal and child undernutrition: Consequences for adult health and human capital' in *Lancet* 371(9609), p340-357.

Victoria, CG., de Onis, M., Hallal, PC., Blossner, M., Shrimpton, R. (2010) 'Worldwide timing of growth faltering: Revisiting implications for interventions' in *Pediatrics* 125(3), e473–e480.

Viljoen, K. (2018) 'Healthy Pregnancy, Healthy Baby: A mobile health service offered in partnership with leading mobile operators in Tanzania' Case Study. GSM Association (GSMA).

World Bank (2017) *World Development Indicators Online.* Washington, DC: The World Bank.

World Health Organization (1997) *WHO Global Database on Child Growth and Malnutrition.* Technical report, World Health Organization.

World Health Organization (2004) *Health System Profile: United Republic of Tanzania*. Technical report, World Health Organization.

World Health Organization (2006) *WHO Child Growth Standards*. Technical report, World Health Organization. URL www.who.int/ childgrowth/standards/Technical_report.pdf.

World Health Organization (2008) *Indicators for Assessing Infant and Young Child Feeding Practices: Part 1 Definitions*. Technical report, World Health Organization.

World Health Organization (2014) *WHO Policy Brief for the Implementation of Intermittent Preventative Treatment of Malaria in Pregnancy Using Sulfadoxine-Pyrimethamine (IPTP-SP)*. Technical report, World Health Organization.

Annex A Community Listing Exercise (CLE) Questionnaire

Study on mobile phone technology based nutrition and health advisory services in Tanzania
COMMUNITY LISTING EXERCISE: Household Questionnaire – August 26, 2016

DRAFT: For Research Purpose only

ENUMERATOR INSTRUCTIONS:

For this survey, use the following definitions of a household, household member and household head:

- ***Household:*** *a group of people who live and eat together, share resources and form a common decision-making unit. The ordinary household is composed of a head of household, his spouse(s), his unmarried children, and possibly his relatives or other persons to whom he is unrelated. The household can be limited to only one person or a person with his children. In a polygamous household where all the spouses do not live in the same concession as their husband, each of the spouses living elsewhere will be listed as a separate household with the persons they live with (the spouse being the head of that household). A tenant who does not take his meals where he lives is considered as a separate household. In a case where a man lives in a concession with his spouse(s) and his children among which some are married, each of the married sons with his spouse(s), his children, and other unmarried dependents under his responsibility be part of his household. In a group of unmarried people living together where everyone has his own means of livelihood, each member of the group will form his own household.*

- ***Household member:*** *anyone who met the criteria for being part of the household more than half of the time during the past 3 months, as well as anyone who recently entered the household through birth or marriage to a household member.*

- ***Household head:*** *the individual who plays a leading role in household decision-making, particularly concerning household economic activity and expenditures. Generally, the person identified by the household as the household head is accepted in this role for the survey.*

Module A: Household Information

ENUMERATOR: Ask the questions below of the household head. If the household head is not available for interview, ask the spouse of the head or the next most responsible household member.

No.	Household Identification	Response	No.	Household Identification	Response
A01	Household Census Identification Number	☐☐☐☐	**A13**	Village (code)	☐☐
A02	Name of the Household Head		**A13a**	Village (name)	☐☐
A02a	Name of respondent (if different from Household Head):		**A14**	Hamlet (code)	☐☐
A03	Does anyone in the household own a mobile phone?	☐ Yes.....1 No.....2	**A15d**	Day of the First visit (dd)	☐☐
A04	How many household members are there in this household?		**A15m**	Month of the First visit (mm)	☐☐
A05	Is there any member of this household that can read and write in Swahili?	☐ Yes.....1 No.....2	**A15y**	Year of first visit (yy)	☐☐
A06	Is any member of this household a pregnant women?	☐ Yes.....1 No.....2	**A16a**	GPS coordinates, Latitude (minutes)	☐☐ ☐☐☐
A06	What is the name of this pregnant woman? *(Enumerator: If more than one, ask for name of pregnant household head or spouse first, then for name of oldest household member who is pregnant.)*		**A16b**	GPS coordinates, Longitude (minutes)	☐☐ ☐☐☐
A07	Is any member of this household under 12 months of age?	☐ Yes.....1 No.....2	**A17**	Code of Interviewer:	☐☐
A08	Does anyone else in this household own a mobile phone?	☐ Yes.....1 No.....2	**A18**	Code of Supervisor:	☐☐
A08a	What is the mobile phone number for this phone? (2 phone numbers)				
A09	Region (code):	☐☐			
A09a	Region (name):				
A10	District (code):	☐☐			
A10a	District (name):				
A11	Division (code):	☐☐			
A11a	Division (name):				
A12	Ward (code):	☐☐			
A12a	Ward leader:				

Annex B Household Survey

Study on mobile phone technology based nutrition and health advisory services in Tanzania
BASELINE SURVEY: Household Questionnaire – October 13, 2016

DRAFT: For Research Purpose only

Outline:

Module A: Household Front Page Identification

Module B: Household Composition and Education
Part 1: Household Roster

Module C: Housing and Assets
Part 1: Housing
Part 2: Current Household Assets

Module D: General Health
Part 1: Health and Vaccinations for those under 5
Part 2: HIV/AIDS Awareness

Module E: Marriage and Fertility History
Part 1: Marriage History (female)
Part 2a: Desired Fertility Preferences (female)
Part 2b: Desired Fertility Preferences (male)

Module F: Mobile phone access and usage
Part 1a: Mobile phone access and usage (female)
Part 1b: Mobile phone access and usage (male)
Part 2: Household mobile usage and access

Module G: Antenatal and Postnatal Care

Module H: IYCF Knowledge and Beliefs
Part 1: IYCF Knowledge and Beliefs (female)
Part 2: IYCF Knowledge and Beliefs (male)

Module I: Nutrition Practices
Part 1: Infant and Young Child Feeding (IYCF) Practices
Part 2: Dietary Diversity

Module J: Trust likelihood of nutrition and health information

Module K: Project Women's Empowerment in Agriculture Index
Part 1: Nutrition and Health Module

Module L: Anthropometry
Part 1: Anthropometry

Universal Codes (Include with all CAPI options):

77=Not applicable
88=Don't know
99=Response refusal

ENUMERATOR INSTRUCTIONS:

For this survey, use the following definitions of a household, household member, household head, primary female respondent, and primary male respondent:

- ***Household:*** *a group of people who live and eat together, share resources and form a common decision-making unit. The ordinary household is composed of a head of household, his spouse(s), his unmarried children, and possibly his relatives or other persons to whom he is unrelated. The household can be limited to only one person or a person with his children. In a polygamous household where all the spouses do not live in the same concession as their husband, each of the spouses living elsewhere will be listed as a separate household with the persons they live with (the spouse being the head of that household) unless the spouses and their families are sharing a meal three or more times a week. A tenant who does not take his meals where he lives is considered as a separate household. In a case where a man lives in a concession with his spouse(s) and his children among which some are married, each of the married sons with his spouse(s), his children, and other unmarried dependents under his responsibility be part of his household. In a group of unmarried people living together where everyone has his own means of livelihood, each member of the group will form his own household.*

- ***Household member:*** *anyone who met the criteria for being part of the household more than half of the time during the past 3 months and is expected to be a member of the household for the next 6 months, as well as anyone who recently entered the household through birth or marriage to a household member. Temporary migrant workers and students residing away from home should be included as household members.*

- ***Household head:*** *the individual who plays a leading role in household decision-making, particularly concerning household economic activity and expenditures. Generally, the person identified by the household as the household head is accepted in this role for the survey.*

- ***Female Respondent:*** *In households with a pregnant woman, the female respondent is the pregnant woman. In households with a child under 12 months of age, the female respondent is the mother of that child. In households where there are two pregnant women or two different women with children under 12 months of age, the spouse of the head of household should be selected as the Female Respondent if she is pregnant or the mother of one of the children under 12 months of age. In the event there are multiple pregnant women or mothers of children under 12 months of age who are also spouses of the head of household, the Female Respondent should be the highest order (earliest) wife. If a woman is both pregnant and has a child under 12 months, then that is pregnant or the mother of the youngest child under 12 months of age.*

- ***Male Respondent:*** *The male respondent should be the spouse or partner of the Female Respondent. If not available, the male head of household can be treated as the Male Respondent. If neither the spouse of the Female Respondent nor the male head of household are available, the father, father-in-law, or brother of the Female Respondent can be selected as the Male Respondent*

Module A: Household Front Page Identification

ENUMERATOR: Ask the questions below of the household head. If the household head is not available for interview, ask the spouse of the head or the next most responsible household member.

No.	Household Identification	Response	No.	Household Identification	Response
A01	Household Identification Number:	☐☐☐☐	**A11a**	Member ID of Primary Respondent:	☐☐
A02	Census number:	☐☐☐	**A11b**	Name of the Primary Respondent:	
A03	Treatment status:	☐☐ Treatment group ..... 1 Control group 2	**A12**	Main language spoken at home:	☐ Swahili ..... 1 Datooga ..... 2 Chaga ..... 3 Makonde ..... 4 Other (specify) 5
A04	Region (code):	☐☐	**A13**	Tribe of household head:	
A04a	Region (name):		**A14**	Location	☐ Urban ..... 1 Rural 2
A05	District (code):	☐☐	**A15**	Religion	Muslim ..... 1 Christian ..... 2 Animist..... 3 Other..... 4
A05a	District (name):		**A16d**	Day of the First visit (dd)	☐☐
A06	Division (code):	☐☐	**A16m**	Month of the First visit (mm)	☐☐
A06a	Division (name):		**A16y**	Year of first visit (yy)	☐☐
A07	Ward (code):	☐☐	**A17d**	Day of second visit (dd):	☐☐
A07a	Ward leader:		**A17m**	Month of second visit (mm):	☐☐
A08	Village (code)	☐☐	**A17y**	Year of second visit (yy)	☐☐

No.	Household Identification	Response	No.	Household Identification	Response
A08a	Village (name)		**A18**	Code of Interviewer:	☐☐
A09	Hamlet (code)	☐☐	**A19**	Code of Supervisor:	☐☐
A10a	Name of the Household Head:		**A20a**	GPS coordinates, Latitude (minutes)	☐☐ ☐☐☐
A10b	Member ID of the Household Head:	☐☐	**A20b**	GPS coordinates, Longitude (minutes)	☐☐ ☐☐☐

CONSENT OF RESPONDENT

Thank you for this opportunity to speak with you. Together with the International Food Policy Research Institute (IFPRI), and Institute for Development Studies (IDS), we are conducting a survey that will provide us with necessary information to carry out research that is designed to help promote the welfare of people in Tanzania; particularly, to improve nutrition and health outcomes. Your household has been chosen by a random selection process.

We value your opinion and there are no wrong answers to the questions we will be asking. We will use approximately 1.5 hours of your time to collect all the information. The survey will collect information about demographics, housing and assets, mobile phone access and usage, consumption, nutrition knowledge and behaviors, and anthropometric data (height and weight) for children. In addition to this, we may access your phone data on use of service from your network provider, for research purposes only.

There will be no cost to you other than your time. There will be no risk as a result of your participating in the study. Your participation in this research is completely voluntary.
You are free to withdraw your consent and discontinue participation in this study at any time and will not be adversely affected. You are also free to refuse to answer any question we may ask.

This study is conducted anonymously. You will only be identified through code numbers. Your identity will not be stored with other information we collect about you. Any information we obtain from you during the research will be kept strictly confidential.

Do not hesitate to ask any questions you may have on the objectives of the study or your participation. Your participation will be highly appreciated. The answers you give will help provide better information to policy-makers, practitioners and program managers so that they can plan for better services that will respond to your needs.

***Enumerator, ask the respondent if they consent to participate in the study*:** "Your participation in this survey is voluntary. You are free to discontinue participation at any time and the investigator will gladly answer any questions that arise during the course of the research. Have you understood the information I gave you and do you agree for you and members of your household to participate in the interview I have just described?"

Contact Persons:

Add a name and contact info from OPM Tanzania

Please tick mark on the right box depending on the respondent's consent
Consent given (household roster, male, and female respondent):

Consent given for anthropometry:

Yes	No

Module B: Household Composition and Education

Module B, Part 1: Household Roster

ENUMERATOR: Ask the questions below of the household head. If the household head is not available for interview, ask the spouse of the head or the next most responsible household member.

Member ID	Name	Sex Male ...1 Female 2	Age (years) If answer is < 5, then → B1_02b If >=5, skip to B1_03	Age in months) If B1_02a<5	Date of birth	Relation to household head	Marital Status (ages 12 and older) If B1_01=2 & B1_04=2|3 Otherwise skip to B1_06a	Is [NAME] pregnant? (ages 15-49 and female) 1. Yes 2. No	Does [NAME]'s husband live in the household? 1. Yes 2. No	Who is NAME's husband? If B1_06a=1	Can [NAME] read and write in Kiswahili? (ages 6 and older)	Can [NAME] read and write in English? (ages 6 and older)	What is the highest level of education [NAME] has completed? (Only for individuals 5 years and older) If B1_07=98, then skip to B1_09	Is [NAME] currently (2016) attending school/ college? (Only for household members age 5-25) Yes ....1 No 2	What is the main activity of [name]? >> skip for those under 12 years of age	In the last 7 days, how many nights has [NAME] stayed at this house?	In the last 7 days, how many days has [NAME] eaten at least one meal with this household?	Does [NAME]'s mother live in the household? 1. Yes 2. No	Who is NAME's biological mother? If B1_13=1	Does [NAME} have a mobile phone? 1. Yes 2. No (ages 12 and over)
ID		Code↑				Code a	Code b				Code c	Code c	Code d		Code e					
MID	**Name**	**B1_01**	**B1_02a**	**B1_02b**	**B1_02c**	**B1_03**	**B1_04**	**B1_05**	**B1_06a**	**B1_06b**	**B1_07a**	**B1_7b**	**B1_08**	**B1_09**	**B1_10**	**B1_11**	**B1_12**	**B1_13**	**B1_14**	**B1_15**

Code a: Relationship with household head

Household head..... 1
Spouse of household head..... 2
Son/daughter 3
Daughter/son -in-law..... 4
Grandson/daughter..... 5
Father/mother 6
Brother/sister 7
Niece/Nephew 8
Household head's cousin 9
Father-in-law/mother-in-law..... 10
Brother/Sister-in-law 11
Spouse's niece/nephew 12
Spouse's cousin..... 13
Primary caregiver.....14
Step-son/Step-daughter.....15
Other (specify) 16

Code b: Marital status code

Unmarried (never married)..... 1
Married, monogamous 2
Married, polygamous 3
Widow/widower..... 4
Divorced 5
Separated/Deserted..... 6

Code c: Literacy

Cannot read and write..... 1
Can sign only..... 2
Can read only 3
Can read and write 4

Code d: Education (Highest class passed)

Never attended school..... 98
Reads in class I 0
Completed class I 1
Completed class II 2

Put number of highest completed class.
For example, if currently in class IV, put 3
(class III completed)

None.....95
Adult School/Madrasa...94
Pre-primary.....0
D1.....1
D2.....2
D3.....3
D4.....4
D5.....5
D6.....6
D7.....7
D8.....8
F1.....9
F2.....10
F3.....11
F4.....12
F5.....13
F6.....14
Diploma.....15
U1.....16
U2.....17
U3.....18
U4.....19
U5.....20

Code e: Activity

Crop Production..... 1
Livestock (including fishing)... 2
Commerce..... 3

Employee 4
Student..... 5
Unpaid housework..... 6
Retired..... 7
Looking for work 8
Other 9
None..... 10

Module C: Housing and Assets

Module C, Part 1: Housing

ENUMERATOR: Ask the questions below of the household head. If the household head is not available for interview, ask the spouse of the head or the next most responsible household member.

C1_01	What is the main material used for the outer walls of the dwelling occupied by your household? [Observation only]	☐	Mud/Mud brick……..1 Wood/Bamboo……...2 Metal/slate/asbestos……3 Stone/Burned bricks……4 Cement/Sandcrete block……5 Landcrete……6 Thatch/ Cardboard……..7 Other (specify)……8
C1_02	What is the main material used for the floor in your dwelling? [Observation only]	☐	Earth/mud/Mud brick……..1 Wood……...2 Stone……3 Cement/ Conrete……4 Ceramic/ tile……5 Other……6
C1_03	What is the main material used for the roof in your dwelling? [Observation only]	☐	Leaves ……...1 Wood……...2 Corrugated metal……3 Cement/ Concrete……4 Asbestos/ Slate/ Tiles……5 Mud brick/Earth……6 Plastic sheeting……..7 Other (specify)……8
C1_04	How many distinct complete rooms does the household occupy, excluding kitchen and bathrooms?		Number
C1_05	What is your main source of drinking water for your household?	☐	Piped into dwelling ..... 1 Public tap ..... 2 Borehole, well with pump ..... 3 Well without pump ..... 4 Spring ..... 5 Pond/Lake/Dam ..... 6 River ..... 7 Rainwater ..... 8 Sachet or bottled water ..... 9 Other (specify) 10

C1_06	What is the main type of toilet used by your household?	☐	Flush / Pour flush Flush to piped sewer system ..... 1 Flush to septic tank ..... 2 Flush to pit (latrine) ..... 3 Flush to somewhere else ..... 4 Flush to unknown place / Not sure / DK where ..... 5 Pit latrine Ventilated Improved Pit latrine (VIP) ..... 6 Pit latrine with slab ..... 7 Pit latrine without slab / Open pit ..... 8 Composting toilet ..... 9 Bucket ..... 10 Hanging toilet, Hanging latrine ..... 11 No facility, Bush, Field ..... 12 Other (*specify*) _____ 13
C1_07	Do you share this toilet facility with other households?	☐	Yes……1 No …….2
C1_08	What is the main type of lighting used by your household?	☐	Electric lights ..... 1 Torch ..... 2 Candles ..... 3 Oil or kerosene lamp ..... 4 Solar lamps ..... 5 Other……...6 None 7
C1_09	What is the main type of cooking fuel used by your household?	☐	Firewood ..... 1 Charcoal ..... 2 Gas/LPG ..... 3 Electricity ..... 4 Other 5

Module C, Part 2: Current Household Assets

C2_ID	Respondent ID		MID

Description of asset	Asset code	Does your household own the item? Yes...1 No....2 >> skip to next row	Quantity
		Code↑	(No.)
Asset	**C2_01**	**C2_02**	**C2_03**
Stove / Gas burner	1		
Metal cooking pots	2		
Bed	3		
Armoire/Cabinet	4		
Table	5		
Chair	6		
Electric fan	7		
Lantern	8		
Iron (charcoal or electric)	9		
Audio cassette/CD player/ Radio	10		
Video/DVD	11		
Sofa	12		
Television (B/W or color)	13		
Refrigerator/Freezer	14		
Sewing machine	15		
Bicycle	16		
Moped	17		
Motor Vehicle	18		
Boat	19		
Saw	20		
Solar energy panel	21		
Plough & yoke for animals	22		
Hoe	23		
Spade/shovel	24		

Description of asset	Asset code	Does your household own the item? Yes...1 No....2 >> skip to next row	Quantity
		Code↑	(No.)
Asset	**C2_01**	**C2_02**	**C2_03**
Reaper/Sickle	25		
Manual sprayer	26		
Rake/Harrow	27		
Wheelbarrow	28		
Carts (hauling)	29		
Tractor	30		
Water pumping set	31		
Fertilizer distributor	32		
Spraying machines (chem./fertilizer)	33		
Harvesting and threshing machine	34		
Bulls/oxen and cows	35		
Sheep and goat	36		
Chicken and other fowl	37		
Donkeys and mules	38		

Module D: General Health

Module D, Part 1: Health and Vaccinations for those under 5

ENUMERATOR: Ask the questions below of the primary female respondent.

Enter the ID of each child born in 2011 or later who is in the household. Ask the questions about all of these children . Begin with the last birth. (***Can prefill on CAPI***)	Name *(Prefill from roster for children under 5)*	How would you rate the general health of [name] [Enumerator: Please read options aloud] Poor….1 Average…2 Good…..3 Very good…4 Excellent….5	Has [name] been ill with a fever at any time in the past 14 days? Yes ….1 No…2	Has [name] been ill with a cough at any time in the past 14 days? Yes…. 1 No…2	Has [name] had diarrhea any time in the past 14 days? Yes ….1 No…2 >> skip to D1_05	How many days did [name] have diarrhea (duration) in the past 14 days?	Did [name] receive the BCG vaccine against tuberculosis, that is, an injection on the right arm or shoulder that usually causes a scar? Yes ….1 No…2 DK	Did [name] receive a polio vaccine, that is, drops in the mouth? Yes ….1 No…2 DK	Did [name] receive the DPT-HB vaccine, that is, an injection given in the thigh or buttocks? Yes ….1 No…2 DK	Did [name] receive a measles or MMR injection, that is, a shot in the arm at the age of 9 months or older – to prevent them from getting measles? Yes ….1 No…2 DK
ID		Code↑	Code↑	Code↑	Code↑	Number	Code↑	Code↑	Code↑	Code↑
MID	**Name**	**D1_01**	**D1_02**	**D1_03**	**D1_04**	**D1_04a**	**D1_05**	**D1_06**	**D1_07**	**D1_08**

Module D, Part 2: HIV/AIDS Awareness (female)

ENUMERATOR: Ask the questions below of the primary female respondent.

D2_ID	Respondent ID		MID

D2_01	Have you ever heard of an illness called AIDS?		Yes ……..1 No……...2 >> skip to next section
D2_02	I don't want to know the results, but were you ever tested for the AIDS virus?		Yes ……..1 No……...2 >> skip to D2_03
D2_02a	I don't want to know the results, but did you get the results of the test?		Yes ……..1 No……...2 >>skip to D2_04 if B1_04=2 or 3
D2_03	I don't want to know the results, but has your partner been tested for the AIDS virus?		Yes ……..1 No……...2 Don't know……..99
D2_04	Have you heard of antiretroviral therapy (ART) *(include local names for ART)*?		Yes ……..1 No……...2 >> skip to next section
D2_05	Do you know where to receive antiretroviral therapy (ART) *(include local names for ART)*?		Yes ……..1 No……...2

Module E: Marriage and Fertility History

Module E, Part 1: Marriage History (female)

ENUMERATOR: Ask the questions below of the primary female respondent.

E1_ID	Respondent ID		MID

E1_01	Are you currently married or living together with a man as if married?		Yes, currently married……..1 >> skip to E1_04 Yes, living with a man……...1 >> skip to E1_04 No, not in union……2
E1_02	Have you ever been married or lived together with a man as if married?		Yes, formerly married……..1 Yes, lived with a man……2 Never been in union……3 >> skip to section E2
E1_03	What is your marital status now? Are you widowed, divorced, or separated?		Widowed……..1 >> skip to E1_09 Divorced……...2 >> skip to E1_09 Separated……3 >> skip to E1_09
E1_04	Is your husband/partner living with you now or is he staying elsewhere?		Living with her……..1 Staying elsewhere……...2
E1_05	Record husband's/partner's member ID from the household roster. If not listed please record '00'.		MID
E1_06	Does your husband/partner have other wives or does he live with other women as if married?		Yes……..1 No……...2 >> skip to E1_09 Don't know……88
E1_07	Including yourself, in total, how many wives or partners does your husband live with now as if married?		Number
E1_08	Are you the first, second…wife?		First.....1 Second.....2 Third.....3 Fourth or higher...4
E1_09	Have you been married or lived with a man only once or more than once?		Only once……..1 More than once.……...2
E1_10a	In what year did you start living with your most recent husband/partner?		Year
E1_10b	In what month did you start living with your most recent husband/partner?		Month >> skip to E2 if E1_09=1
E1_11a	Now I would like to ask you about your first husband/partner. In what year did you start living with him?		Year
E1_11b	In what month did you start living with your first husband/partner?		Month

Module E, Part 2a: Desired Fertility Preferences (female)

ENUMERATOR: Ask the questions below of the primary female respondent.

E2a_01	Are you currently pregnant?	☐	Yes……..1 No …….2 >> skip to E2a_05 Not sure……3 >> skip to E2a_05
E2a_01a	How far along in the pregnancy are you?(Enumerator: Encourage respondent to respond in number of weeks. If she is unable to then prompt her to respond in months.)		Time
E2a_01a_unit	Time unit	☐	Months……1 Weeks……..2
E2a_01b	How long ago did you find out you were pregnant? (Enumerator: Encourage respondent to respond in number of weeks. If she is unable to then prompt her to respond in months.)		Time
E2a_01b_unit	Time unit	☐	Months……1 Weeks……..2
E2a_02	After the child you are expecting now, would you like to have another child, or would you prefer to not have any more children?	☐	Have another child……...1 No more……2 >> skip to E2a_04 Undecided……3 >> skip to E2a_04
E2a_03	After the birth of the child you are expecting now, how long would you like to wait before trying to conceive your next child?		Time (in months or years)
E2a_04	Do you think you will use a contraceptive method to delay or avoid pregnancy at any time in the future?		Yes……...1 >> skip to E2a_09 No……2 >> skip to E2a_09 Undecided……3 >> skip to E2a_09
E2a_05	Would you like to have a/another child, or would you prefer to not have any/more children?	☐	Have another child……...1 No more……2 >> skip to E2a_07 Undecided……3 >> skip to E2a_07
E2a_06	How long would you like to wait before trying to conceive your next child?		Time (in months or years)
E2a_07	Are you (individual) using a contraceptive method?	☐	Yes……...1 No……2 >>skip tp E2a_09
E2a_08	What method of contraception are you using? [Multiple answers possible]	☐	IUD……...1 Injectables……2 Pill……...3 Condom….4 Female Condom…5 Foam/Jelly…6 Sterilization…7 Other…8
E2a_09	How many children do you currently have?		Number >> skip to E2a_10 if E2a_05=2
E2a_09a	How many additional children would you like to have?		Number

E2a_10	Among the children that you (still) wish to have, how many would you like to be boys, girls, and for how many would the sex not matter? Boys?		Number
E2a_11	Girls?		Number
E2a_12	Sex would not matter?		Number
E2a_13	Does your partner want the same number of children that you want, or do they want more or fewer than you want?	☐	Same number……...1 More children……2 Fewer children……...3

Module E, Part 2b: Desired Fertility Preferences (male)

ENUMERATOR: Ask the questions below of the primary male respondent.

E2b_ID	Respondent ID	☐	MID Not present….98 >> skip to moudle F

E2b_01	Is your wife (partner)/ Any of your wives (partners) currently pregnant?	☐	Yes……..1 No ……2 >> skip to E2b_05 Not sure……..3 >> skip to E2b_05
E2b_01a	How far along in the pregnancy is your wife (partner)		Time
E2b_01a_unit	Time unit	☐	Months……1 Weeks……..2
E2b_01b	How long ago did you find out your wife (partner) was pregnant?		Time
E2b_01b_unit	Time unit	☐	Months……1 Weeks……..2
E2b_02	After the child you are expecting now, would you like to have another child, or would you prefer to not have any more children?	☐	Have another child……...1 No more……2 >> skip to E2b_04 Undecided……3 >> skip to E2b_04
E2b_03	After the birth of the child you are expecting now, how long would you like to wait before you conceive your next child?		Time (in months and years)
E2b_04	Do you think you will use a contraceptive method to delay or avoid pregnancy at any time in the future?		Yes……...1 >> skip to E2b_09 No……2 >> skip to E2b_09 Undecided……3 >> skip to E2b_09
E2b_05	Would you like to have a/another child, or would you prefer to not have any/more children?	☐	Have another child……...1 No more……2 >> skip to E2b_07 Undecided……3 >> skip to E2b_07
E2b_06	How long would you like to wait before you conceive your next child?		Time (in months and years)

E2b_07	Are you using a contraceptive method?	☐	Yes……..1 No……2 >>skip tp E2b_09
E2b_08	What method of contraception are you (individual) using?[Multiple answers possible]	☐	IUD……..1 Injectables……2 Pill……..3 Condom….4 Female Condom…5 Foam/Jelly…6 Other…7
E2b_09	How many total living children do you have?		Number >> skip to E2b_10 if E2b_05=2
E2b_09a	How many additional children would you like to have?		Number
E2b_10	Among the children that you (still) wish to have, how many would you like to be boys, girls, and for how many would the sex not matter? Boys?		Number
E2b_11	Girls?		Number
E2b_12	Sex not matter?		Number
E2b_13	Does your partner want the same number of children that you want, or do they want more or fewer than you want?	☐	Same number……..1 More children……2 Fewer children……..3

Module F: Mobile phone access and usage

Module F, Part 1a: Mobile phone access and usage (female)

ENUMERATOR: Ask the questions below of the primary female respondent.

F1a_ID	Respondent ID		MID

F1a_00	Do you own a working mobile phone?		Yes……..1 >> skip to F1a_01b No……...2
F1a_01	Do you have access to a working mobile phone?		Yes……..1 No……...2 >> skip to next section
F1a_01a	Who owns the mobile phone you have access to?		Spouse …….1 Another member in this household…….2 >>Skip to F1a_01c
F1a_01b	Please list the mobile phone numbers of phones you own. *(Enumerator: enter 999999999 if the respondent cannot recall the phone number. Enter all numbers they own))*		Mobile phone numbers
F1a_01c	Please list the mobile phone number you have access to. *(Enumerator: enter 999999999 if the respondent cannot recall the phone number.)*		Mobile phone numbers
F1a_02a	**[Enumerator: Is the phone with the respondent?]**		1 Yes 2 No >> skip to F1a_03
F1a_02b	**[Enumerator: Is the phone on or can it be switched on?]**		1 Yes 2 No >> skip to F1a_03
F1a_02c	**[Enumerator: Dial and verify the number. If it does not ring, ask the respondent to check the number and keep trying until you hear the phone ring]** Was the number verified?		1 Yes 2 No [CAPI should give a warning before proceeding to next question]

F1a_03	Who else has access to this mobile phone? *(Multiple answers allowed)*		Spouse …….1 Another member in this household……2 A neighbour……....3 A family member in the village……..……4 A friend in the village……5 Another person outside the village……..6 Other……7
F1a_04 _a	Did you use your mobile phone in the last fourteen days to make calls?		1 Yes 2 No
F1a_04 _b	Did you use your mobile phone in the last fourteen days to receive calls?		1 Yes 2 No
F1a_04 _c	Did you use your mobile phone in the last fourteen days to write text messages?		1 Yes 2 No
F1a_04 _d	Did you use your mobile phone in the last fourteen days to receive text messages?		1 Yes 2 No
F1a_04 _e	Did you use your mobile phone in the last fourteen days to send mobile money?		1 Yes 2 No
F1a_04 _f	Did you use the mobile phone in the last fourteen days to receive mobile money?		1 Yes 2 No
F1a_04 _g	Did you use the mobile phone in the last fourteen days to use mobile internet (e.g., Whatsapp, Facebook, and email).		1 Yes 2 No
F1a_05	How many times did you use the phone in the last two weeks? (Read options aloud)		Once or twice.....1 On some days.....2 Most days.....3 A few times every day.....4 Multiple times every day...5 >> skip to F1a_08 if F1a_04_c=2 & F1a_04_d=2
F1a_06	Does anyone in your household know how to send and receive text messages?		Yes……..1 No……...2 >> skip to F1a_13 Skip if F1a_04_c or F1a_04_d=1
F1a_07	Has anyone in your household ever sent a text message?		Yes……..1 No……...2
F1a_08	Did you use your mobile phone to get health advice of any kind?		Yes……..1 No……...2
F1a_09	Have you received automatic text messages about with information about healthy food consumption during pregnancy, breastfeeding information or complementary food information at any time in the past two years?		Yes……..1 No……...2 skip to F1a_13
F1a_10	Who is sending these messages?		Government health facility (dispensary, health center, hospital) 1 NGO……2 Religious institution……3 Other……96

			Don't Know……98
F1a_11	When was the last time you received this message?	☐	Yesterday……..1 Last week……..2 Two weeks ago……3 Last month……..4 Two months ago…….5 Less than six months ago but more than two months ago…6 In the last year but more than six months ago….7 13 to 24 months ago……8

F1a_13	How much does you spend in total on airtime on all phone in an average month?		Tanzanian shillings
F1a_14	Which is your main phone number?		
F1a_15	Can you receive signal for this network at your household compound? **[Enumerator: Ask this of each mobile phone in the roster]**	☐	Yes……..1 No……...2
F1a_16	Can you receive signal for this network somewhere in the village? **[Enumerator: Ask this of each mobile phone in the roster]**	☐	Yes……..1 No……...2
F1a_17	How long does it take to walk from your household to the nearest place where you can receive a signal from this network? **[Enumerator: Ask this of each mobile phone in the roster]**		Minutes
F1a_18	Do you typically charge your phone at home?	☐	1. Yes >> skip to next section 2. No
F1a_19	How long does it take to get to the nearest place where you typically charge your phone?	☐	1. Less than 10 minutes 2. 10-30 minutes 3. 31 minutes to an hour 4. More than an hour
F1a_20	How much do you typically pay in a month for charging your phone?		Tanzanian shillings

Module F, Part 1b: Mobile phone access and usage (male)

ENUMERATOR: Ask the questions below of the primary male respondent.

F1b_ID	Respondent ID		MID

F1b_00	Do you own a mobile phone?		Yes……..1 >> skip to F1b_02 No……...2
F1b_01	Do you have access to a mobile phone?		Yes……..1 No……..2 >> skip to next section
F1b_01a	Who owns the mobile phone you have access to?		Spouse …….1 Another member in this household…….2 >>Skip to F1b_01c
F1a_01b	Please list the mobile phone numbers of phones you own. *(Enumerator: enter 999999999 if the respondent cannot recall the phone number. Enter all numbers they own)*		Mobile phone number
F1b_01c	Please list the mobile phone number you have access to. *(Enumerator: enter 999999999 if the respondent cannot recall the phone number. Enter all numbers they have access to)*		Mobile phone number
F1b_02a	**[Enumerator: Is the phone with the respondent?]**		1 Yes 2 No >> skip to F1b_03
F1b_02b	**[Enumerator: Is the phone on or can it be switched on?]**		1 Yes 2 No >> skip to F1b_03
F1b_02c	**[Enumerator: Dial and verify the number. If it does not ring, ask the respondent to check the number and keep trying until you hear the phone ring]** Was the number verified?		1 Yes 2 No [CAPI should give a warning before proceeding to next question]
F1b_03	Who else has access to this mobile phone? *(Multiple answers allowed)*		Spouse …….1 Another member in this household…….2 A neighbour……....3 A family member in the village……..…….4 A friend in the village……5 Another person outside the village……6 Other……7
F1b_04_a	Did you use the mobile phone in the last fourteen days to make calls?		1 Yes 2 No

F1b_04 _b	Did you use the mobile phone in the last fourteen days to receive calls?	☐	1 Yes 2 No
F1b_04 _c	Did you use the mobile phone in the last fourteen days to write text messages?	☐	1 Yes 2 No
F1b_04 _d	Did you use the mobile phone in the last fourteen days to receive text messages?	☐	1 Yes 2 No
F1b_04 _e	Did you use the mobile phone in the last fourteen days to send mobile money?	☐	1 Yes 2 No
F1b_04 _f	Did you use the mobile phone in the last fourteen days to receive mobile money?	☐	1 Yes 2 No
F1b_04 _g	Did you use the mobile phone in the last fourteen days to use mobile internet (e.g., Whatsapp, Facebook, and email).	☐	1 Yes 2 No
F1b_05	How many times did you use the phone in the last two weeks? (Read options aloud)	☐	Never……..1 1-4……2 5-9……3 10-15……4 >15 ……..5 >> skip to F1b_08 if F1b_04_c=2 & F1b_04_d=2
F1b_06	Does anyone in your household know how to send and receive text messages?		Yes……..1 No……...2 >> skip to next section
F1b_07	Has anyone in your household ever sent a text message?		Yes……..1 No……...2
F1b_08	Did you use your mobile phone to get health advice of any kind?	☐	Yes……..1 No……...2
F1b_09	Have you received automatic text messages about with information about healthy food consumption during pregnancy, breastfeeding information or complementary food information at any time in the past two years?	☐	Yes……..1 No……...2 skip to F1b_13
F1b_10	Who is sending these messages?		Government health facility (dispensary, health center, hospital) 1 NGO/International Organization ……2 Religious institution (church, mosque)…..3 Other……96 Don't Know……98
F1b_11	When was the last time you received this message?	☐	Yesterday……..1 Last week……..2 Two weeks ago……3 Last month……..4 Two months ago……5 Less than six months ago but more than two months ago…6 In the last year but more than six months ago….7 13 to 24 months ago……8

F1b_13	How much do you spend in total on airtime on your phone in an average month?		Tanzanian shillings
F1b_14	Which is your main phone number?		
F1b_15	Can you receive signal for this network at your household compound? **[Enumerator: Ask this of each mobile phone in the roster]**	☐	Yes……..1 No……...2
F1b_16	Can you receive signal for this network somewhere in the village? **[Enumerator: Ask this of each mobile phone in the roster]**	☐	Yes……..1 No……...2
F1b_17	How long does it take to walk from your household to the nearest place where you can receive a signal from this network? **[Enumerator: Ask this of each mobile phone in the roster]**		Minutes
F1a_18	Do you typically charge your phone at home?	☐	1. Yes >> skip to next section 2. No
F1b_19	How long does it take to get to the nearest place where you typically charge your phone/s?	☐	1. Less than 10 minutes 2. 10-30 minutes 3. 31 minutes to an hour 4. More than an hour
F1b_20	How much do you typically pay in a month for charging your phone/s?		Tanzanian shillings

Module G: Antenatal and Postnatal Care

ENUMERATOR: Ask the questions below of the primary female respondent.

G1_ID	Respondent ID		MID
G1_00	One or more births in 2013 or later?		Yes……1 No births in 2013 or later……2 >> skip to next section

Enter the ID of each birth in 2013 or later. Ask the questions about all of these births. Begin with the last birth. (*Can prefill on CAPI)*	Is this child living or dead? 1. Living 2. Died since birth 3. Died at birth (Skip this question if child comes from roster)	When you got pregnant with [NAME] did you see anyone for antenatal care? Yes...1 No....2 >> skip to G1_08	Where did you receive antenatal care for this pregnancy? Home….1 Govt health facility (dispensary, health center, hospital)….2 Private hospital/clinic……..3 Religious/voluntary hospital…….4 Other……6	How many months pregnant were you when you first received antenatal care?	How many times did you receive antenatal care?	As a part of your antenatal care during this pregnancy, were any of the following done at least once: a. Was your blood pressure measured? b. Did you give a urine sample? c. Did you give a blood sample? Yes...1 No....2			During this pregnancy, were you told about the signs of pregnancy complications ? Yes...1 No....2	During this pregnancy, were you given an injection in the arm to prevent the baby from getting tetanus, that is, convulsions after birth? Yes...1 No....2
G1_01	**G1_02**	**G1_03**	**G1_04**	**G1_05**	**G1_06**	**G1_07a**	**G1_07 b**	**G1_07c**	**G1_08**	**G1_09**

Enter the ID of each birth in 2013 or later. Ask the questions about all of these births. Begin with the last birth. (*Can prefill on CAPI)*	During this pregnancy, were you given or did you buy any iron syrup/iron or iron/folate tablets (tablets to increase blood count)? Yes...1 No....2	During the whole pregnancy how many days did you take the tablets or syrup?	During this pregnancy, did you have difficulty with your vision during the daylight? Yes...1 No....2	During this pregnancy, did you suffer from night blindness? Yes...1 No....2	During this pregnancy, did you sleep under a bed net? Yes...1 No....2	During this pregnancy, did you take any drugs to keep you from getting malaria? Yes...1 No....2	Who assisted with the delivery? Doctor/AMO….1 Clinical Officer….2 Nurse/midwife……..3 Village health worker…..4 Trained TBA/TBA …….5	Where did you give birth? Home….1 Govt health facility (dispensary, health center, hospital) ….2 Private hospital…..3 Religious/ voluntary hospital….4	After the birth of this child were you instructed to attend a follow up visit at your nearest dispensary? 1.Yes 2. No	Did you have a post-natal follow up visit within the first week after [name]'s birth? 1.Yes 2. No	How many post-natal follow up visits did you have within the first 28 days after [name]'s birth?	How many post-natal follow up visits did you have within the first 42 days after [name]'s birth?	Were you visited in your home by a community health worker during the first few days after the delivery of [name]? 1. Yes 2. No

							Other……6	Other……6					
G1_01	**G1_10**	**G1_11**	**G1_12**	**G1_13**	**G1_14a**	**G1_14b**	**G1_15**	**G1_16**	**G1_17**	**G1_18**	**G1_19**	**G1_20**	**G1_21**

Module H: IYCF Knowledge and Beliefs

Module H, Part 1: IYCF Knowledge and Beliefs (female)

Note: The following questions relate to primary female's knowledge of feeding practices, not their actual behavior, which may or may not be consistent with their awareness and knowledge due to a number of circumstances. Enumerator read to respondent: "Now I would like to ask you a few questions on what you think are the best ways to feed a baby. I am interested in what you personally think, not what others may tell you, or what you are able to provide to a baby given the circumstances."

ENUMERATOR: Ask the questions below of the primary female respondent.

Question number	Question	Response	Response option
MEM ID	Copy the respondent's ID from module B	☐	Member ID
H1_01	How long after birth should a baby start breastfeeding?	☐	Immediately ..... 1 Less than 1 hour after delivery ..... 2 Some hours later but less than 24 hrs .... 3 1 day later ..... 4 More than 1 day later ..... 5 Do not think baby should be breastfed ... 6 Don't know 88
H1_02	What should a mother do with the "first milk" or colostrums?	☐	Throw it away and start breastfeeding when the real milk comes in ..... 1 Give it to her baby by breastfeeding soon after birth ..... 2 Other (specify) ..... 3 Don't know 88
H1_03	How often should a baby breastfeed? *(Allow multiple options)*	☐	Whenever baby wants ..... 1 When you see the baby is hungry ..... 2 When the baby cries ..... 3 Other (specify) ..... 4 Don't know 88
H1_04	If a mother thinks her baby is not getting enough breast milk, what should she do? *(Allow multiple options)*	☐	Breastfeed more often/more frequently .. 1 Give other liquids/foods ..... 2 Mother needs to drink more water ..... 3 Mother needs to eat more food ..... 4 Other (specify) ..... 5 Don't know 88
H1_05	Do you think that infants under 6 months of age should be given water if the weather is very hot?	☐	Yes ..... 1 No ..... 2 Don't know 88
H1_06	At what age should a baby first start to receive liquids (including water) other than breast milk?		Months Don't know 88

Question number	Question	Response	Response option
H1_07	At what age should a baby first start to receive foods in addition to breast milk?		Months Don't know....88
H1_08	Name one thing that can happen to children if they do not get enough iron (either in their diet or via iron supplements). (*Allow multiple responses)*	☐	Impaired learning....1 Impaired development....2 Lower height....3 Weakened immune defense....4 Feel tired....5 Become anemic....6 Other (specify)....7 Don't know....88
H1_09	What food are rich in Vitamin A? (*Allow multiple responses)*	☐	Orange fruits/vegetables....1 Green leaves....2 Eggs....3 Liver....4 Breast milk....5 Cow milk....6 Meat/fish....7 Other (specify)....8 Don't know....88
H1_10	What foods are rich in iron?	☐	Green leafy vegetables....1 Liver....2 Meat/fish....3 Beans....4 Other (specify)....8 Don't know....88
H1_11	What are the best strategies to protect a child from intestinal worms? (*Allow multiple responses)*	☐	Wash hands (soap not mentioned)....1 Wash hands with soap....2 Clip nails....3 Wear pants....4 Wash fruits and vegetables....5 Wear shoes....6 Give treated water....7 Other (specify)....8 Don't know....88

Module H, Part 2: IYCF Knowledge and Beliefs (male)

Note: The following questions relate to primary female's husbands knowledge of feeding practices, not their actual behavior, which may or may not be consistent with their awareness and knowledge due to a number of circumstances.

ENUMERATOR: Ask the questions below of the primary male respondent.

Question number	Question	Response	Response option
MEM ID	Copy the respondent's ID from module B	☐	Member ID
H2_01	How long after birth should a baby start breastfeeding?	☐	Immediately ..... 1 Less than 1 hour after delivery ..... 2 Some hours later but less than 24 hrs .... 3 1 day later ..... 4 More than 1 day later ..... 5 Do not think baby should be breastfed ... 6 Don't know 88
H2_02	What should a mother do with the "first milk" or colostrums?	☐	Throw it away and start breastfeeding when the real milk comes in ..... 1 Give it to her baby by breastfeeding soon after birth ..... 2 Other (specify) ..... 3 Don't know 88
H2_03	How often should a baby breastfeed? *(Allow multiple options)*	☐	Whenever baby wants ..... 1 When you see the baby is hungry ..... 2 When the baby cries ..... 3 Other (specify) ..... 4 Don't know 88
H2_04	If a mother thinks her baby is not getting enough breast milk, what should she do? *(Allow multiple options)*	☐	Breastfeed more often/more frequently .. 1 Give other liquids/foods ..... 2 Mother needs to drink more water ..... 3 Mother needs to eat more food ..... 4 Other (specify) ..... 5 Don't know 88
H2_05	Do you think that infants under 6 months of age should be given water if the weather is very hot?	☐	Yes ..... 1 No ..... 2 Don't know 88
H2_06	At what age should a baby first start to receive liquids (including water) other than breast milk?		Months Don't know 88
H2_07	At what age should a baby first start to receive foods in addition to breast milk?		Months Don't know 88

Question number	Question	Response	Response option
H2_08	What can happen to children if they do not get enough iron, either in their diet or via iron supplements?. (*Allow multiple responses)*	☐	Impaired learning....1 Impaired development....2 Lower height....3 Weakened immune defense....4 Feel tired....5 Become anemic....6 Other (specify)....7 Don't know....88
H2_09	What food are rich in Vitamin A? (*Allow multiple responses)*	☐	Orange fruits/vegetables....1 Green leaves....2 Eggs....3 Liver....4 Breast milk....5 Cow milk....6 Meat/fish....7 Other (specify)....8 Don't know....88
H2_10	What are the best strategies to protect a child from intestinal worms? (*Allow multiple responses)*	☐	Wash hands (soap not mentioned)....1 Wash hands with soap....2 Clip nails....3 Wear pants....4 Wash fruits and vegetables....5 Wear shoes....6 Give treated water....7 Other (specify)....8 Don't know....88

Module I: Nutrition Practices

Module I, Part 1: Infant and Young Child Feeding (IYCF) Practices

Instructions: Ask the following for the child aged 0 to 12 months and the next youngest child (under 5 years) in the household. Respondent should be the mother of the index child. ***[CAPI should repeat section]***

ENUMERATOR: Ask the questions below of the primary female respondent.

Question Number	Questions		Code
Member ID- CHILD	Copy the child's ID from Module B	Mem ID ☐	Mem ID
Member ID- MOTHER	Note the child's mother's ID from Module B	Mem ID ☐	Mem ID
I1_00	Is the father part of the household roster?	☐	1. Yes 2. No
Mem ID- FATHER	Note the child's father's ID from Module B	Mem ID ☐	Mem ID
I1_01_d	What is the day of birth of this child?	☐	DD
I1_01_m	What is the month of birth of this child?	☐	MM
I1_01_y	What is the year of birth of this child?	☐	YY
I1_03	Did anyone help you put the baby to the breast after birth?	☐	Yes ..... 1 No..... 0
I1_03a	How soon after birth did you put the baby to breast?	☐	Immediately ..... 1 Less than 1 hour after delivery ..... 2 Some hours later but less than 24 hrs.... 3 1 day later..... 4 More than 1 day later..... 5
I1_04a	During the first 3 days after birth, was <NAME> given anything to eat or drink other than breastmilk?	☐	Yes …..1 No…..2 → Skip to I1_05 DK…….88 → Skip to I1_05

Question Number	Questions		Code
I1_04b	What was the child given? *(Multiple response possible)*	☐ ☐ ☐	Honey ..... 1 Plain water ..... 2 Sugar/glucose water ..... 3 Tea ..... 4 Cow's milk ..... 5 Other (specify) ..... 6 Do not remember 88
I1_05	Did you give the child colostrum?	☐	Yes (gave to child) ..... 1 No (did not give to child) ..... 2 DK ……..88
Now we would like to ask you about what the child is eating now:			
I1_06	Is the child still breastfeeding?	☐	Yes ..... 1**>>I1_09** No ..... 2 Never ..... 3**>>I1_09**
I1_07	If no, at what age did you stop breastfeeding the child?	☐	Month Don't Know/cannot remember 88
I1_08	Why did you stop breastfeeding? *(Do not prompt)*	☐	Problems with breast (pain) ..... 1 Child not suck well ..... 2 Not enough time to feed child ..... 3 Child already grown up/ No need for breast feeding ..... 4 Mother got pregnant ..... 5 New baby born ..... 6 Cracked nipples ..... 7 Felt not enough breast milk ..... 8 Other (specify) 9
I1_09	At what age did you start giving the following liquids/foods to the child? *Note: If mother fed her child any of the following food within the first 29 days (less than 1 months of age), this can be noted as "0" month.*		
	1. Water	☐	Month At "0" month of age ..... 0 At "1" month of age ..... 1 At "2" months of age ..... 2 At "3" months of age ..... 3 .... At "12" months of age ..... 12 So on Don't Know ..... 88 Not given yet 98
	2. Cow/Goat milk	☐	
	3. Other non breast milk liquids (sugar/glucose water, tea, fruit juice etc.)	☐	
	4. Rice gruel, etc.	☐	
	5. Semi-solid foods (soft rice, mashed potato, ripe banana, other mashed family foods, uji etc.)	☐	
	6. Solid foods (such as rice, wheat, etc.)	☐	
	7. Fish (including daga)	☐	
	8. Meat (chicken, beef, goat etc.)	☐	
	9. Eggs	☐	
	10. Legumes (pulse, peas, etc)	☐	

Question Number	Questions		Code
	11. Green vegetables	☐	
	12. Snack foods	☐	
The following questions are based on previous day recall, i.e., the last 24 hours during the day and night.			
I1_10	How many times did you breastfeed [NAME] yesterday, during the day or night?	☐	Number of times Stopped breast feeding/Never breast fed 98
I1_11	Other than breast milk, how many times did [NAME] drink other milk, formula or yogurt yesterday, during the day or night? *Note: Do not include number of times the child was breastfed in this question. This variable is only to capture milk or milk products **other than breast milk**.*	☐	Number of times Not given yet..... 98
I1_12	How many times did [NAME] eat solid, semi-solid or soft foods other than liquids yesterday, during the day or night? ***Note: Semi-solid** foods such as soft rice, mashed potato, ripe banana, other mashed family foods etc. **Solid** foods such as rice, wheat, puffed/pressed rice etc. Meals include both meals and snacks (other than trivial amounts)*	☐	Number of times Not given yet..... 98
I1_13	Yesterday (during the day or the night) did you give any of the following liquids to the child?		
		Yesterday 1. Yes 2. No → skip to next row	
I1_13a	Breast milk	☐	Yes ..... 1 No..... 2
I1_13b	Plain water	☐	
I1_13c	Baby formula (prepared food for child)	☐	
I1_13d	Any other kind of milk (powder, cow/goat milk etc.)	☐	
I1_13e	Fruit juice or juice drinks	☐	
I1_13f	Clear broth	☐	
I1_13g	Water-based liquids, teas, sugar water, coffee	☐	
I1_13h	Thin porridge	☐	
I1_13i	Yogurt	☐	
I1_14	Yesterday (during the day and the night), did you use a baby bottle to feed the child?	☐	Yes ..... 1 No..... 2

Question Number	Questions		Code
I1_15	Did your child eat (or drink) any of the following foods yesterday (during the day or night)?	Yesterday 1. Yes 2. No → skip to next row	
I1_15a	Bread, rice, noodles, porridge or other foods made from grains		
I1_15b	Pumpkin, red or yellow yams, carrots, sweet potatoes that are yellow or orange inside		
I1_15c	White potatoes, cassava, or any other foods made from roots, tubers or plantain		
I1_15d	Any dark green, leafy vegetables (e.g. amaranth, cow pea leaves, sweet potato leaves, spinach, cassava leaves)		
I1_15e	Ripe mangoes, papaya, melon,		
I1_15f	Any other fruits or vegetables (e.g. bananas, avocados, tomatoes, oranges, apples)		
I1_15g	Any meat, such as beef, pork, lamb, goat, chicken, or duck		
I1_15h	Eggs		
I1_15i	Fresh or dried fish or shellfish (e.g. prawn, lobster)		
I1_15j	Any foods made from beans, peas, lentils, nuts, or seeds		
I1_15k	Yogurt, cheese, or other milk products		
I1_15l	Any oil, fats, or butter, or foods made with any of these		
I1_15m	Any sugary foods such as chocolates, sweets, candies, pastries, cakes, or biscuits		
I1_15n	Condiments for flavor, such as peppers, spices, herbs or fish powder		
I1_15o	Plumpy'nut, Nutributter		
I1_15p	Foods made with palm oil, or oil fortified with Vitamin A		

Module I, Part 2: Dietary Diversity

Instructions: Ask the following for the primary female respondent. For the women's diet questions include food consumed at home or outside the home. For the household level, exclude foods purchased and eaten outside the home.

Question Number	Questions	Female Respondent (15-49 years)		Code
Member ID-	Copy the respondent's ID from Module B	Mem ID ☐		Mem ID
I2_01_d	What is the day of birth?	☐		DD
I2_01_m	What is the month of birth?	☐		MM
I2_01_y	What is the year of birth?	☐		YY
The following questions are based on previous day recall, i.e., Yesterday during the day and the night.				
I2_02	**Enumerator: *Ask each food item for the primary female respondent and for the household separately.*** *Female*: Yesterday (during the day or the night) did you eat or drink any [food item]? Please include food consumed at home or outside of the home. *Household*: Yesterday (during the day or the night) did anyone in your household eat or drink any [food item]? Please only include food consumed at home.	Female Respondent	Household	
I2_02a	Milk such as tinned, powdered, or fresh animal milk	☐	☐	Yes .....1 No 2
I2_02b	Tea or coffee	☐	☐	
I2_02c	Any other liquids (juice, cocoa)	☐	☐	
I2_02d	Bread, rice, noodles, porridge or other foods made from grains	☐	☐	
I2_02e	Pumpkin, red or yellow yams, carrots, sweet potatoes that are yellow or orange inside	☐	☐	
I2_02f	White potatoes, cassava or any other foods made from roots, tubers or plantain	☐	☐	
I2_02g	Any dark green, leafy vegetables (e.g. amaranth, cow pea leaves, sweet potato leaves)	☐	☐	
I2_02h	Ripe mangoes, papaya, melon, nere	☐	☐	
I2_02i	Any other fruits or vegetables (e.g. bananas, avocados, tomatoes, oranges, apples)	☐	☐	
I2_02j	Liver, kidney, heart or other organ meats	☐	☐	
I2_02k	Any meat, such as beef, pork, lamb, goat, chicken, or duck	☐	☐	
I2_02l	Eggs	☐	☐	
I2_02m	Fresh or dried fish or shellfish (e.g. prawn, lobster)	☐	☐	
I2_02n	Any foods made from beans, peas, lentils, nuts, or seeds	☐	☐	

Question Number	Questions	Female Respondent (15-49 years)		Code
I2_02o	Yogurt, cheese, or other milk products			
I2_02p	Any oil, fats, or butter, or foods made with any of these			
I2_02q	Any sugary foods such as chocolates, sweets, candies, pastries, cakes, or biscuits			
I2_02r	Condiments for flavor, such as peppers, spices, herbs or fish powder			
I2_02s	Plumpy'nut, Nutributter			
I2_02t	Foods made with palm oil, or oil fortified with Vitamin A			

Module J: Trust likelihood of nutrition and health information

ENUMERATOR: Ask the questions below of the primary female respondent.

Question number	Question	Response	Response option	
MEM ID	Copy the respondent's ID from module B		Member ID	
J1_00	Is [source] a source of information on nutrition and health for you?		Spouse .....1 Other family .....2 Friend/neighbor.....3 Automated SMS (from an NGO, private organization, religious or voluntary organization, government) .....4 Government health worker (at a health facility or elsewhere)…5 Community Health Worker from my village 6	Clinic/hospital(run by private organization, religious or voluntary organization, NGO,)7 TV/Radio/Poster (Other public advertisement) ..... 8 Traditional health worker ..... 9 Health worker from my village (non-government) ..... 10 Other (specify) 11
J1_00a	Which source is most important? (Up to 2 options)		Spouse .....1 Other family .....2 Friend/neighbor.....3 Automated SMS (from an NGO, private organization, religious or voluntary organization, government) .....4 Government health worker (at a health facility or elsewhere)…5 Community Health Worker from my village 6	Clinic/hospital(run by private organization, religious or voluntary organization, NGO,)7 TV/Radio/Poster (Other public advertisement) ..... 8 Traditional health worker ..... 9 Health worker from my village (non-government) ..... 10 Other (specify) 11

Instructions: Below are a series of statements that you may agree or disagree with. Using the scales below indicate your agreement with each item. Please be open and honest in your response.

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Response Code
J1_01a	If I were to receive any breastfeeding, complementary feeding, or general health information from my spouse I would feel confident and trust it completely.	1	2	3	4	5	
J1_02a	If I were to receive any breastfeeding, complementary feeding, or general health information from my other family members I would feel confident and trust it completely.	1	2	3	4	5	
J1_03a	If I were to receive any breastfeeding, complementary feeding, or general health information from my neighbors and friends I would feel confident and trust it completely.	1	2	3	4	5	
J1_04a	If I were to receive any breastfeeding, complementary feeding, or general health information from TV/radio/posters and other public advertisements I would feel confident and trust it completely.	1	2	3	4	5	
J1_05a	If I were to receive any breastfeeding, complementary feeding, or general health information from automated SMS messages from an NGO, private organization, religious, or voluntary organization or the government, I would feel confident and trust it completely.	1	2	3	4	5	
J1_06a	If I were to receive any breastfeeding, complementary feeding, or general health information from a clinic or hospital run by a private organization, religious or voluntary organization, NGO or government, I would feel confident and trust it completely.	1	2	3	4	5	
J1_07a	If I were to receive any breastfeeding, complementary feeding, or general health information from my community health worker I would feel confident and trust it completely.	1	2	3	4	5	
J1_08a	If I were to receive any breastfeeding, complementary feeding, or general health information from the health worker from an NGO or project I would feel confident and trust it completely.	1	2	3	4	5	
J1_09a	If I were to receive any breastfeeding, complementary feeding, or general health information from a traditional health worker I would feel confident and trust it completely.	1	2	3	4	5	
J1_10a	If I were to receive any breastfeeding, complementary feeding, or general health information from the government health worker (at a facility or elsewhere) I would feel confident and trust it completely.	1	2	3	4	5	

Module K: Anthropometry

Module K, Part 1: Anthropometry –

Note: Measure child ≤12 months old. If the mother of the child is not a household member (for example, mother is deceased) put 98

ENUMERATOR: The child measured should be the son or daughter of the primary female respondent.

Member ID	Name	Mother's MID	Child's date of birth					Height (cm)	Weight (kg)	How was this child's height measured?	If not measured, why? (Mark height and weight '00' if not measured)
			Day (Prepopulate from Module I)	Month (Prepopulate from Module I)	Year (Prepopulate from Module I)	How old was (name) at his last birthday? (Years) *(Note: Compare with K1_04 and correct if necessary)*	Confirmed with birth certificate or health card? Yes.....1 No 2			Lying down...1 Standing up...2 (*Note: All children under 24 months old must be measured in lying position*)	Have measured.....1 Absent.....2 Sick.....3 Refused to give measurement .....4 Other (specify) 5
MID	**Name**	**K1_01**	**K1_02**	**K1_03**	**K1_04**	**K1_04b**	**K1_05**	**K1_06**	**K1_07**	**K1_08**	**K1_09**

L. Closing and additional consent if treatment

Enumerator read aloud: "Thank you for your participation in the study. We are now done with the interview. Your household may be selected to receive free text messages that contain health and nutrition information that would benefit you and your children. This service is called mNutrition' provided by Wazazi Nipendeni. Are you willing to give consent for your household to receive these text messages?"

ENUMERATOR: Ask the questions below of the primary female respondent and the primary male respondent.

Question Number	Questions		Code
L1_01	Does the primary female respondent consent to receiving the text messages?	☐	Yes 1 No 2
L1_02	Does the male respondent consent to receiving the text messages? (skip if male respondent does not own a phone)	☐	Yes 1 No 2

Annex C IFPRI IRB Approval

Date: September 13, 2016

IRB application approval number: 16-9-13 (Temporary)
IRB #00007490
FWA #00005121

Study Project Title: External Evaluation of Mobile Phone Technology Based Nutrition Advisory Services in Africa
Division: PHND
PI: Daniel Gilligan
Country of study: Tanzania

Date of IRB approval: 09/13/2016
Date of Expiration: 09/12/2017

Dear Dr. Gilligan,

Your application to conduct the study entitled, External Evaluation of Mobile Phone Technology Based Nutrition Advisory Services in Africa, has been reviewed and approved by IFPRI's Institutional Review Board. The study meets the criteria for expedited review using survey procedures as set forth in the code of federal regulations (45 CFR 46.110 Category 7) and presents no more than minimal risks to human subjects. Proper consent requirements have also been met. The IRB has taken note that an award is pending and therefore has assigned a temporary application approval number to this study. When the project has been funded please provide the project number immediately so that an updated approval letter can be issued with the new project information and so that our files can be complete.

This approval is for the period of one year. If you wish to continue this study beyond that time you must submit an application to continue along with the instruments/documentation 6 weeks in advance of the expiration date listed above. Should **any changes become necessary (i.e procedures, methodologies) or be made or added to this study, you must immediately notify the IRB. No activity should commence without IRB modification approval.**

As a reminder the IRB requires that all staff directly working with human subjects in research complete IFPRI'S CITI ethics training course. This letter indicates that the project complies with the IFPRI IRB's ethical guidelines. In cases where local approval is needed, it is the responsibility of the researcher to obtain this approval and comply with local guidelines. Please keep the IRB advised of this.

We wish you all the best in your research efforts. If you have any questions please do not hesitate to contact Olivette Burton, IFPRI IRB Coordinator via phone or the email address copied on this correspondence.

Sincerely,

Eduardo Maruyama
IRB Chair
IFPRI-IRB@cgiar.org

Annex D COSTECH IRB Approval

TANZANIA COMMISSION FOR SCIENCE AND TECHNOLOGY
(COSTECH)

A. I Hassan Mwinyi Road
P.O. Box 4302
Dar es Salaam
Tanzania

RESEARCH PERMIT

No. 2016-310-NA-2016-139

25th July 2016

1. Name : Daniel Orth Gilligan
2. Nationality : American
3. Title : External Evaluation of Mobile Phone Technology based Nutrition Advisory Services in Africa
4. Research shall be confined to the following region(s): Arusha, Dodoma, Iringa, Shinyanga, Tanga
5. Permit validity from: 25th July 2016 to 24th July 2017
6. Contact/Collaborator: Mr. Deogardius Medardi, OPM Tanzania, P.O. Box 33296, Dar es Salaam
7. Researcher is required to submit progress report on quarterly basis and submit all Publications made after research.

M. Mushi
for: DIRECTOR GENERAL

TANZANIA COMMISSION FOR SCIENCE AND TECHNOLOGY
(COSTECH)

Telephones: (255 - 022) 2775155 - 6, 2700745/6
Director General: (255 - 022) 2700750&2775513
Fax: (255 - 022) 2775313

Email: relearance@costech.or.tz
In reply please quote: **CST/RCA 2016/139**

Al' Hassan Mwinyi Road
P.O. Box 4302
Dar es Salaam
Tanzania

25th July 2016

Director of Immigration Services
Ministry of Home Affairs
P.O. Box 512
DAR ES SALAAM

Dear Sir/Madam,

RESEARCH PERMIT

We wish to introduce **Daniel Orth Gilligan** from **USA** who has been granted Research Permit **No. 2016–510-NA-2016-139** dated **25th July 2016**

The permit allows him/her to do research in the country "**External Evaluation of Mobile Phone Technology based Nutrition Advisory Services in Africa**"

We would like to support the application of the researcher(s) for the appropriate immigration status to enable the scholar(s) begin research as soon as possible.

By copy of this letter, we are requesting regional authorities and other relevant institutions to accord the researcher(s) all the necessary assistance. Similarly the designated local contact is requested to assist the researcher(s).

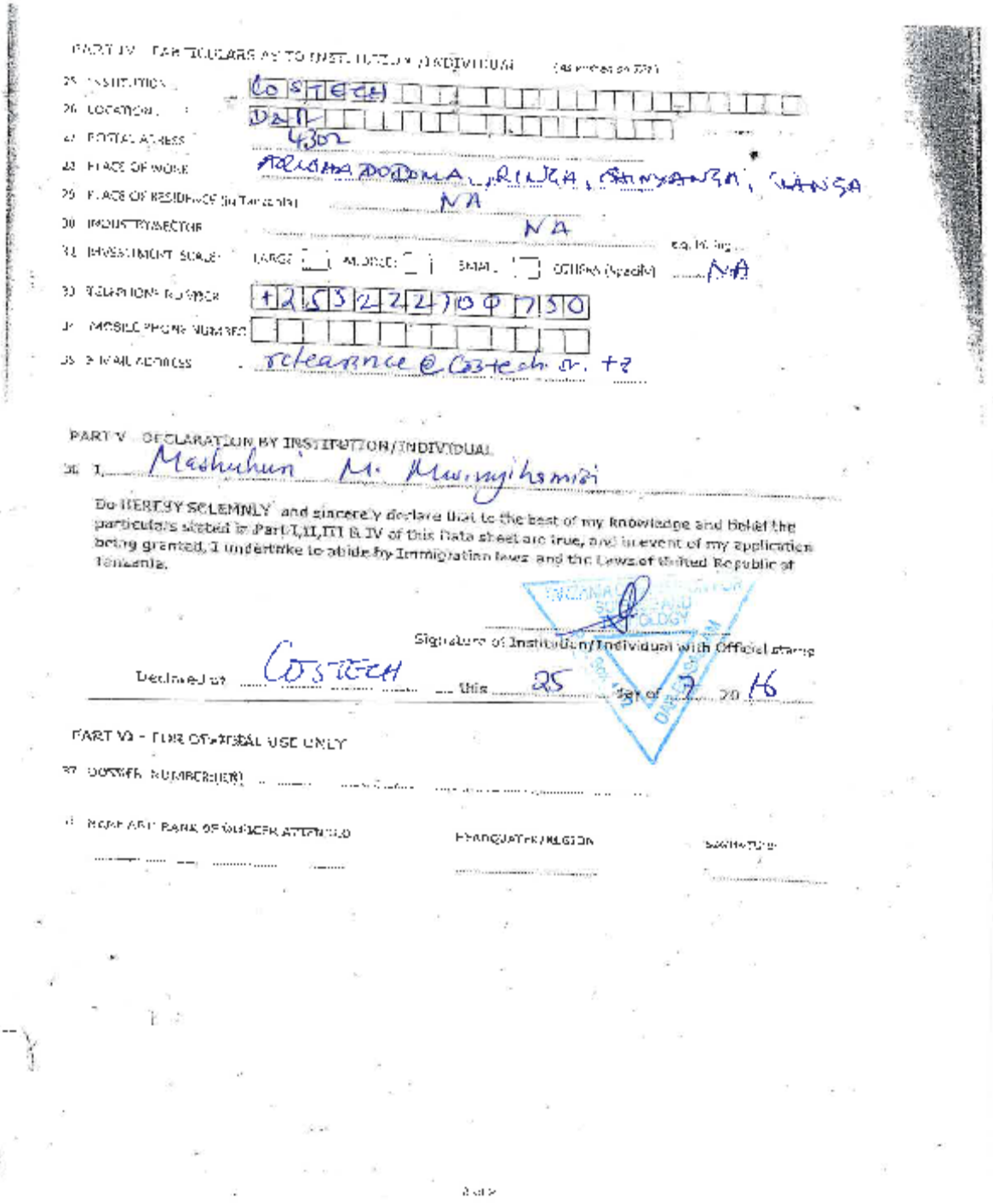
Yours faithfully

M. MOSS

For: DIRECTOR GENERAL

CC:

1. Regional Administrative Secretary: **Arusha, Dodoma, Iringa, Shinyanga, Tanga**
2. Local Contact: **Mr. Deogardios Mediardi, OPM Tanzania. P.O. Box 33296, Dar es Salaam**
3. Co-Researchers: **None**



PART IV - PARTICULARS AS TO INSTITUTION/INDIVIDUAL (as appears on P.I.)

25 INSTITUTION: COSTECH

26 LOCATION: DAR

27 POSTAL ADDRESS: 4302

28 PLACE OF WORK: MOROGORO, DODOMA, MWANZA, MBEYA, KANSA

29 PLACE OF RESIDENCE (in Tanzania): NA

30 INDUSTRY/SECTOR: NA

31 INVESTMENT SCALE: LARGE [ ] MEDIUM [ ] SMALL [ ] OTHERS (specify) NA

32 TELEPHONE NUMBER: +255 22 2700750

33 MOBILE PHONE NUMBER: [ ]

34 E-MAIL ADDRESS: rclearance@costech.or.tz

PART V - DECLARATION BY INSTITUTION/INDIVIDUAL

35 I, Mashuhuri M. Mwingihamisi

DO HEREBY SOLEMNLY and sincerely declare that to the best of my knowledge and belief the particulars stated in Part I, II, III & IV of this Data sheet are true, and in event of my application being granted, I undertake to abide by Immigration laws and the Laws of United Republic of Tanzania.

Signature of Institution/Individual with Official stamp

Declared at COSTECH this 25 day of 7 20 16

PART VI - FOR OFFICIAL USE ONLY

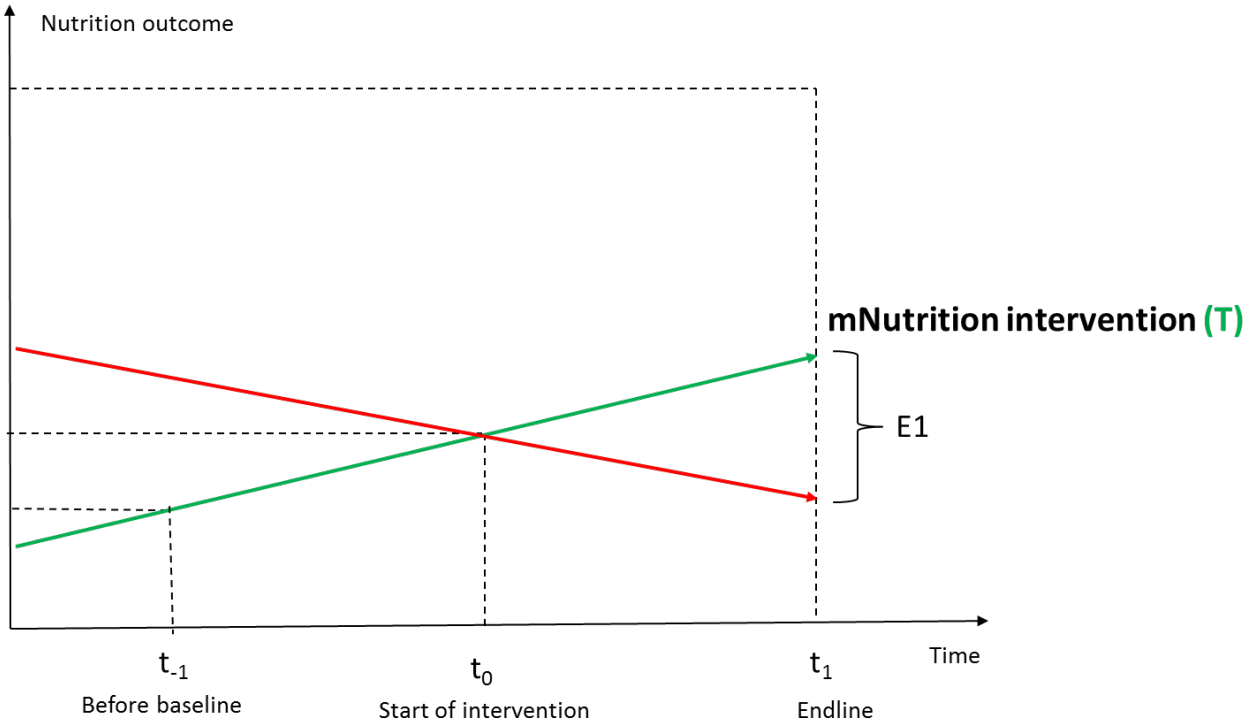
37 DOSSIER NUMBER(S): [ ]

38 NAME AND RANK OF OFFICER ATTENDED: [ ] FRANQUATER/REGION: [ ] SIGNATURE: [ ]

Annex E Cluster Randomized Controlled Trial Details

As with any thoughtful attempt to estimate the causal effect of an intervention, the quantitative evaluation must resolve the fundamental problem of causal inference (Holland, 1986): researchers can never observe the value of an outcome for the same individual or household after having been both exposed to the intervention and not having been exposed to the intervention at the same point in time. Inference regarding the effect of a programme—in this case the treatment is being offered access to the mNutrition programme on a mobile phone—is typically based on comparing outcomes between a group of beneficiary individuals or households who are exposed to the programme and a group of non-beneficiary individuals or households who are not exposed to the programme. However, without the ability to affect which individuals and households are exposed to the treatment, a simple comparison of outcomes by beneficiary status is likely to confound the true causal effect of the programme with selection bias, the effect of observable and unobservable differences between the group that chooses to participate and the group that does not. shows one hypothetical example of how selection bias can lead to misguided conclusions about the impact of an intervention on a nutrition outcome.

Figure 9 Possible Selection Bias in Impact Evaluations

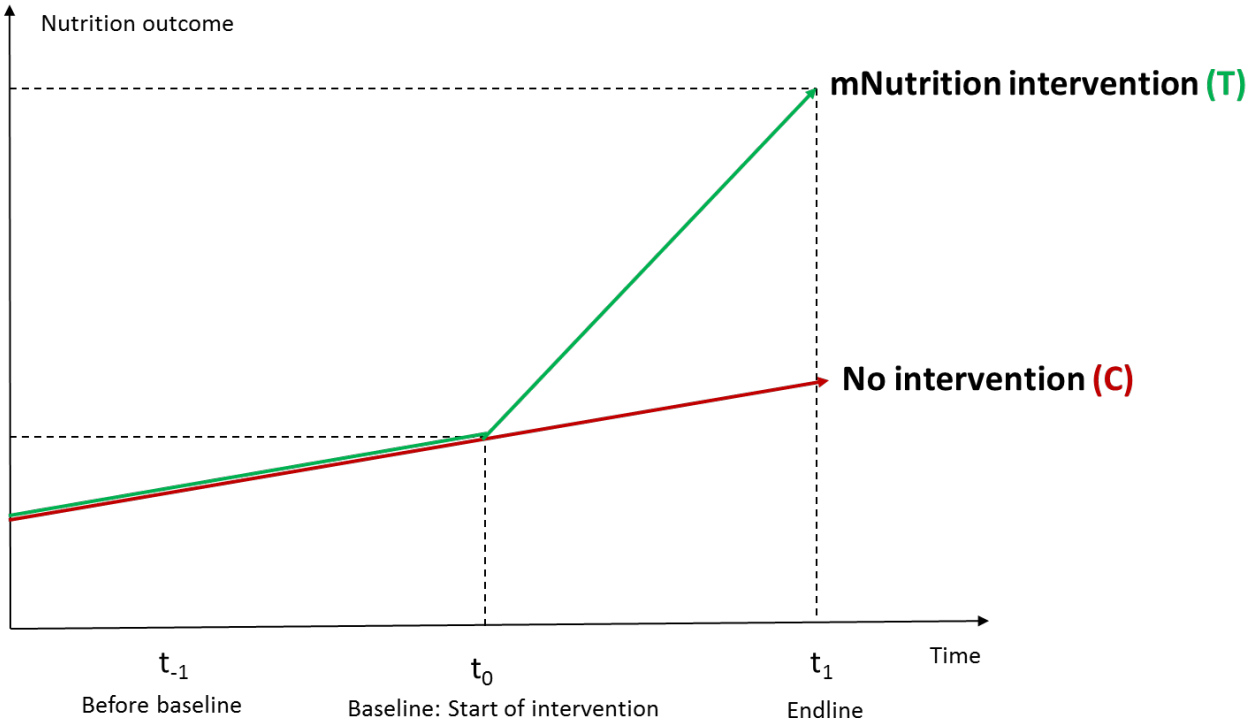


In Figure 7.1, the mean values for a nutrition outcome are plotted separately for a group that self-selects into receiving the mNutrition intervention (T) and a group that elects not to receive the intervention (C). The trends for the nutrition outcome are not parallel to one another, but the mean nutrition outcome for the treatment group is equal to the mean nutrition outcome for the control group at the time of the start of the intervention (t_0). Estimating a simple difference between the average nutrition outcome in the treatment group and the average nutrition outcome in the control group at endline, t_1, or a difference-in-differences treatment effect using the observed values of the outcome at time t_0 would identify a positive treatment effect equal to E1. However, given the drastically different time trends in the nutrition outcome between the two groups it is unlikely that E1 represents the true causal treatment effect of participation in the programme. Instead, the pre-intervention non-parallel trends in the outcome suggest that other observed or unobserved differences between the two groups are more probable causes of the endline difference in outcomes.

To ensure that the evaluation accurately measures the causal impact of mNutrition, the quantitative evaluation is based on a carefully designed cluster randomized controlled trial (cRCT). Among the

randomly selected sample of villages participating in the evaluation, IFPRI randomly assigned households in half of the villages to a treatment group—where sampled households receive a door-to-door offer of access to the mNutrition content—and households in the other half of participating villages to a control group—where the sampled households received no similar offer. Because the villages assigned to the treatment and control groups were randomly selected, any average difference in outcomes between households in treatment and control villages at endline can be attributed to their access to the mNutrition programme. Compared to the scenario in Figure 7.1, where potential beneficiaries non-randomly select into programme participation, the random assignment of villages to treatment groups guarantees that the units allocated to each group will have similar values for the outcomes of interest at the time the intervention and during any pre-intervention period. Figure 7.2 displays how the effect of the mNutrition intervention on a nutrition outcome could be estimated, in the context of an evaluation with baseline and endline data collection.

Figure 10 Estimation of Programme Effects



The villages assigned to the treatment group and those assigned to the control group have identical values of the nutrition outcome on average until the start of the intervention, both in levels at time t_0 and in trends between t_{-1} and t_0. Outcomes begin to diverge as the mNutrition programme begins and the villages allocated to the treatment group start to be exposed to the treatment. Calculating the difference in the nutrition outcome at time t_1 between the treatment and control groups now gives an unbiased estimate of the impact of mNutrition.[50]

In addition to the village level randomization, the evaluation also includes a second stage household level randomization within treatment villages. Specifically, households in treatment villages where both the mother of the young child or pregnant woman (the primary female) and the primary male own distinct mobile phones will be randomly allocated to either just receive the mNutrition content on the mobile phone of the primary female (T-F), or to receive the mNutrition content on the mobile phone of the primary female and the mobile phone of the primary male (T-F+M). By comparing behaviours and outcomes between treatment households with two distinct mobile phones who were randomly assigned to receive the mNutrition content on both phones and those that were not, learning will be generated about how

[50] Given that there is no difference in levels between the two groups when the intervention begins and the baseline survey is conducted (t_0), the differences-in-differences treatment effect would be identical to the simple differences estimate discussed above.

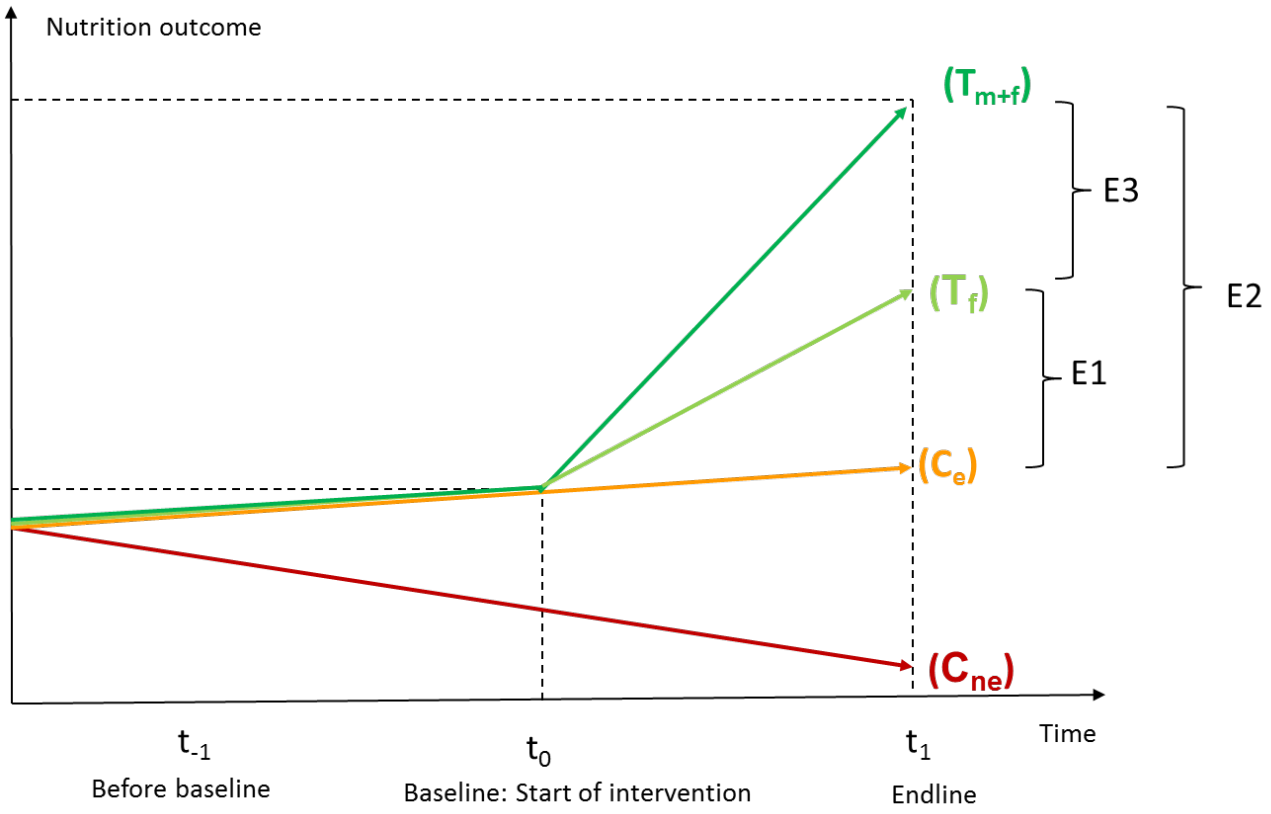
information flows between spouses. Figure 7.3 illustrates how this second stage household level randomization will be used to calculate a different set off treatment effects.

The four lines represent the paths of the nutrition outcome for four distinct groups. The line labelled C_{ne} shows the control group trend for the nutrition outcome in households that were not eligible for the second stage household level randomization—those households where the primary male and primary female do not both own distinct mobile phones. The C_e line plots the nutrition outcome trend for households in the control group that met the criteria of having a primary male and primary female who both own mobile phones. The T_f line shows the outcome trend for households in the treatment group that were eligible for the second stage randomization and assigned to just receive the mNutrition content on the mobile phone of the primary female. The T_{m+f} line shows the outcome trend for households in the treatment group that were eligible for the second stage randomization and assigned to receive the mNutrition content on the mobile phones of both the primary male and primary female.

Figure 7.3 illustrates several important features of the household level randomization. First, in order to estimate whether sending the mNutrition programme to the phone of the primary male, in addition to the phone of the primary female, has a differential impact on nutrition outcomes, it is not necessary for the households eligible for the second stage randomization to be observably or unobservably similar to non-eligible households; that the outcome trend for the C_{ne} group is distinct and non-parallel to the outcome trends for the other three groups does not make it more difficult to identify a causal answer to this research question. Instead, we need it to be true that the households in the treatment group that are eligible for the household randomization and that are allocated to the T-F+M group would have the same level of the nutrition outcome at time t_1 in the absence of the intervention as households in the treatment group that are eligible for the household randomization and that are allocated to the T-F group. Because both groups are required to meet the same set of eligibility criteria and their assignment to the T-F or T-F+M group is entirely dictated by the random allocation mechanism controlled by the research team, this is likely to be satisfied. We can then calculate the effect of sending additional messages to the mobile phone of the primary male by estimating the difference between the average nutrition outcome among households in the T-F+M group (line T_{m+f}) and the average nutrition outcome among households in the T-F group (line T_f) at time t_1; the treatment effect is represented by quantity E3 in Figure 7.3.[51]

[51] Alternatively, we could estimate the impact of the additional messages sent to the mobile of the primary male in three steps. First calculate E2, the difference between the mean nutrition outcome among those in the treatment group eligible for the household randomization who are allocated to the T-F+C group (line T_{m+f}) and the mean nutrition outcome among those in the control group that would have been eligible for the household randomization had they been in the treatment group (line C_e). Second, estimate E1, the difference between the mean nutrition outcome among those in the treatment group eligible for the household randomization who are allocated to the T-F group (line T_f) and the mean nutrition outcome in the control group households that would have been eligible for the household randomization had they been in the treatment group. Third, calculate E3 by subtracting E1 from E2; E3 represents the causal effect of sending the mNutrition messages to the mobile of the primary male, in addition to sending them to the mobile of the primary female.

Figure 11 Second Stage Household Level Randomization in Treatment Villages

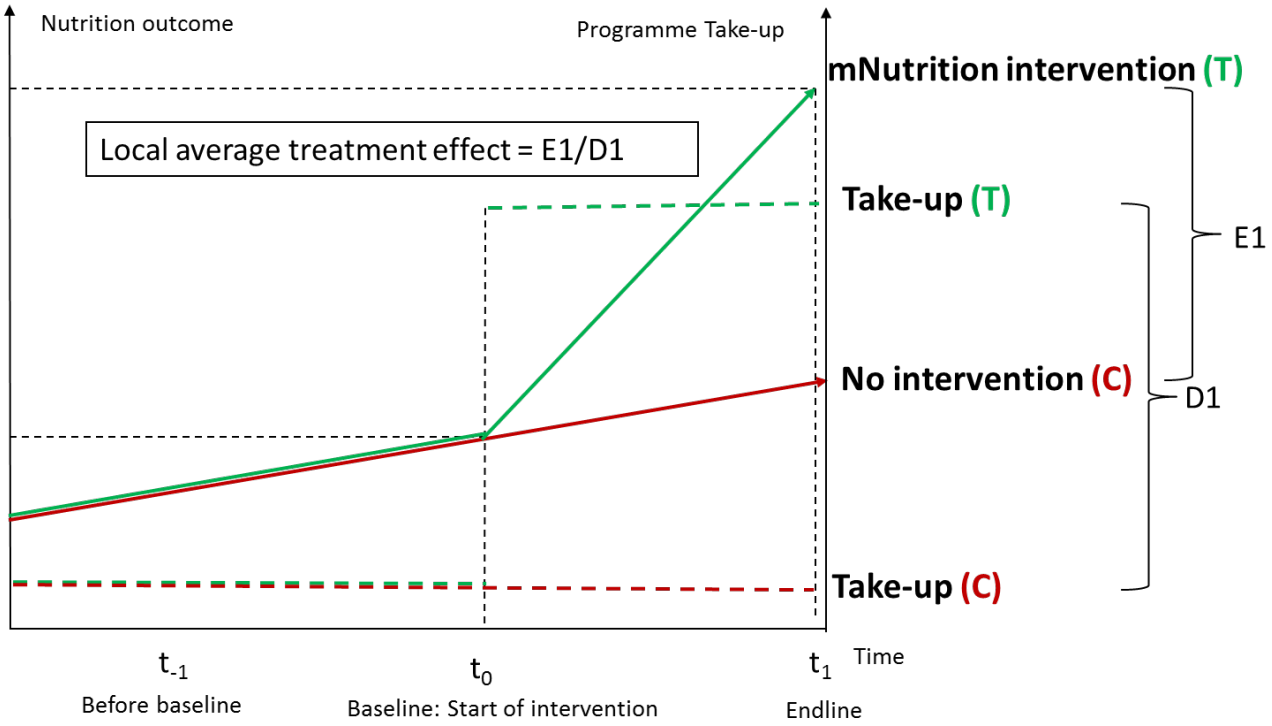


Implicitly assumed in the description above is that all households in treatment villages agree to receive the mNutrition content on their mobile phones and none of the households in control villages register their mobile phones to receive mNutrition content during the study period. In practice, this assumption is unlikely to be satisfied. Some treatment households will decline to provide consent to receive the mNutrition programme and some control households will register for the content. This imperfect compliance with treatment assignment, while important to account for, is unlikely to affect whether the evaluation can identify the causal effect of the programme on outcomes. The impact of being offered access to the mNutrition content can still be calculated by comparing average outcomes at endline between households in treatment villages and households in control villages. However, the impact of registering for the mNutrition content on a mobile phone for those households induced to take-up the programme by the offer can also be identified by scaling the average difference in the outcome across the treatment and control villages by the average difference in registration rates for the mNutrition content across the treatment and control villages.

displays visually the distinction between these two programme impacts. The average difference in the nutrition outcome is represented by the quantity E1. This is identical to the difference in mean outcomes as shown in Figure 7.2. However, Figure 7.4 now also shows the average difference in take-up, or registration, for the mNutrition programme between the treatment and control villages (quantity D1 in Figure 7.4). If quantity D1 is less than one, indicating that there is imperfect compliance with the treatment assignment, then we can also calculate what is known as the local average treatment effect (LATE) for households that comply with their treatment assignment.[52] The LATE of participating in the mNutrition programme for these household compliers is simply the ratio of the difference in the average nutrition outcome between treatment and control households (E1) to the difference in the average take-up of the mNutrition programme between treatment and control households (D1).

[52] Compliers are the set of households that would participate in the programme if they were assigned to the treatment group and that would not participate in the programme if they were assigned to the control group.

Figure 12 Observed Programme Participation and Outcomes at Endline



To better elucidate how the two different randomizations—village- and household-level—map to our analysis plans for research questions 1) through 5), we list the research questions below by which randomization will be the basis for inference.

Village-Level Randomization (groups T and C):

1. What is the impact of the mNutrition programme on women's dietary diversity?

2. What is the impact of the mNutrition programme on Infant and Young Child Feeding (IYCF) practices?

3. What is the impact of the mNutrition programme on nutritional status for children under twelve months of age at baseline?

4. What is the impact of the mNutrition programme on nutrition knowledge, including knowledge of IYCF practices, among pregnant women and the caregivers of young children?

Household-Level Randomization (groups T-F and T-F+M):

5. Does sending the mNutrition programme to the mobile phones of both the primary female—either the pregnant woman or the mother of the child under twelve months of age at baseline—and the primary male—typically the spouse of the primary female—have a differential impact on the other primary and secondary outcomes?

Therefore, our analyses for research questions 1) through 4) will be based on the comparison for the relevant outcome between the T and C groups, while for research question 5) inference will be based on the random allocation of eligible households in treatment villages to either the T-F or T-F+M sub-groups.

As discussed above, the random allocation of villages to the treatment or control group and the random allocation of households in treatment villages to the T-F group or the T-F+M group should ensure that comparing outcomes across any two groups (treatment with control or T-F with T-F+M) yields unbiased estimates of the causal effect of the programme. While the theoretical basis for estimating causal impacts through a randomized experiment is clear, the data collected through the baseline survey offer an

opportunity to empirically assess whether the random assignment of treatment was successful. By comparing baseline measures of the primary and secondary outcomes—as well as other observable characteristics likely to be strongly correlated with the outcomes—we can empirically test whether it is likely that treatment status will be uncorrelated with other factors that may influence outcomes at endline. Though not a perfect test for the presence of bias at endline, the finding of few differences between households assigned to the different treatment groups at baseline sends a strong signal that the evaluation is likely to be able to estimate the unbiased causal effects of the programme.

Annex F Sample Differences and Balance Tests for the Household Level Randomization

7.3 Comparing Households Eligible for the Household Randomization to the Overall Sample

7.3.1 Household Demographics, Physical Structure, Amenities, and Wealth

We begin by showing means and standard deviations for household demographic characteristics, the physical structure of the household, amenities, and measures of household wealth in Table 7.1. The first column presents the number of observations with a non-missing value in the full sample, the second column presents the mean and standard deviation (below in parentheses) for the characteristic in the full sample, the third column displays the number of households that were eligible for the household level randomization and had a non-missing value, the fourth column presents the mean and standard deviation for the characteristic in the sample that was eligible for the household randomization, and the fifth column shows the p-value from a t-test of the null hypothesis of no difference between the means for the two samples.

Of the total number of treatment households, 276 households were identified as being eligible for the household level randomization, just under 20% of the total number. Eligible households are observably different from the full sample. They are significantly less likely to be female headed, 2.9% of eligible households as compared to 12.9% of the full sample, more likely to have a head of household with some formal education (98.9% compared to 93.0%), more likely to have a head of household that is married and monogamous (84.4% compared to 74.2%), more likely to have a primary female that is married and monogamous (74.6% compared to 67.7%), and they have progress out of poverty scores that are significantly higher on average (PPI score of 42.6 compared to 39.3). These differences indicate that the sub-sample eligible for the household randomization may be significantly wealthier than the overall sample. Ex-ante, it is not necessarily clear how these differences would affect the return to participation in the mNutrition programme. On the one hand, wealthier, better educated households may find it easier to understand the information in the SMS messages or they may be more likely to have the resources necessary to act on some of the messaging; conversely, it is possible that wealthier households may already be aware of the information contained in the mNutrition programme, and their nutrition related behaviours may therefore already be compliant with the courses of action suggested by the programme.

Table 7.1: Demographics, by mNutrition household randomization eligibility status

	N	All	N	Eligible for Household Randomization	P-value
Household size	2,833	5.276	276	5.380	0.701
		(2.110)		(2.300)	
Female-Headed Household	2,833	0.129	276	0.029	0.000
		(0.335)		(0.168)	
Age of Household Head	2,833	38.608	276	37.514	0.054
		(13.475)		(11.763)	
Household Head has some education	2,833	0.930	276	0.989	0.000
		(0.254)		(0.104)	
Household Head: Never Married	2,833	0.028	276	0.004	0.000
		(0.164)		(0.060)	
Household Head: Married, Monogamous	2,833	0.742	276	0.844	0.000
		(0.437)		(0.363)	
Household Head: Married, Polygamous	2,833	0.139	276	0.138	0.964

	N	All	N	Eligible for Household Randomization	P-value
		(0.346)		(0.345)	
Household Head's main activity is crop production	2,833	0.731	276	0.736	0.978
		(0.443)		(0.442)	
Age of the Primary Female	2,833	26.999	276	26.931	0.556
		(6.622)		(6.231)	
Primary Female has some education	2,833	0.936	276	0.960	0.279
		(0.245)		(0.196)	
Primary Female: Never Married	2,833	0.208	276	0.167	0.038
		(0.406)		(0.373)	
Primary Female: Married, Monogamous	2,833	0.677	276	0.746	0.005
		(0.468)		(0.436)	
Primary Female: Married, Polygamous	2,833	0.089	276	0.080	0.577
		(0.284)		(0.271)	
Order of wife - first	281	0.306	23	0.217	0.308
		(0.462)		(0.422)	
Order of wife - second	281	0.598	23	0.609	0.789
		(0.491)		(0.499)	
Order of wife - third	281	0.075	23	0.087	0.906
		(0.263)		(0.288)	
Order of wife - fourth	281	0.021	23	0.087	0.231
		(0.145)		(0.288)	
Duration of primary female living with most recent husband/partner (months)	2,099	88.763	224	87.790	0.492
		(73.083)		(73.620)	
Primary Female's main activity is crop production	2,833	0.768	276	0.728	0.078
		(0.422)		(0.446)	
Female focus child (under 12 months)	1,928	0.512	162	0.481	0.736
		(0.500)		(0.501)	
Percent of school-aged children going to school	2,629	49.877	260	50.679	0.815
		(40.054)		(40.367)	
Total PPI score	2,833	39.337	276	42.554	0.000
		(14.146)		(14.838)	
Consumer Durable Asset Index	2,833	0.000	276	0.722	0.000
		(2.065)		(2.239)	
Household Production Asset Index	2,833	-0.000	276	0.467	0.000
		(1.428)		(1.841)	
Household Livestock Asset Index	2,833	-0.000	276	0.039	0.017
		(1.388)		(0.935)	
Household Total Asset Index	2,833	-0.000	276	0.855	0.000
		(2.263)		(2.545)	
Household has access to improved sanitation	2,833	0.757	276	0.793	0.168
		(0.429)		(0.406)	
Household has access to improved drinking water sources	2,833	0.824	276	0.797	0.738
		(0.381)		(0.403)	

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. To be in the sub-set households had to reside in a treatment village and the primary female and the primary male had to each own a distinct mobile phone. Standard deviations are in parentheses.

7.3.2 Primary Outcomes and Child Anthropometry

Table 7.2 continues to assess the comparability between the two samples by checking for differences in primary outcomes and child anthropometry. One caveat to the reported output is that for a number of the tests, a lack of statistical power is an issue. There are just 154 observations with non-missing child anthropometry data in the sub-sample eligible for the household randomization, and the sample size is even smaller for some of the IYCF practices measures—fewer than 20 for whether children aged 12-15 months are still fed breast milk.

Table 7.2: Primary Outcomes, by mNutrition household randomization eligibility status

	N	All	N	Eligible for Household Randomization	P-value
Child age in months (under 12 months)	1,928	5.521	162	5.300	0.547
		(3.421)		(3.441)	
Male child (under 12 months)	1,928	0.488	162	0.519	0.736
		(0.500)		(0.501)	
Child height	1,856	61.845	154	61.850	0.893
		(6.306)		(6.310)	
Child weight	1,857	6.595	154	6.594	0.817
		(1.653)		(1.648)	
Child height-for-age Z-score	1,852	-1.425	154	-1.326	0.363
		(1.190)		(1.275)	
Child weight-for-age Z-score	1,855	-0.670	154	-0.621	0.696
		(1.182)		(1.207)	
Child weight-for-height Z-score	1,851	0.450	154	0.437	0.823
		(1.209)		(1.139)	
Child stunted	1,852	0.296	154	0.266	0.447
		(0.457)		(0.443)	
Child underweight	1,855	0.121	154	0.136	0.302
		(0.327)		(0.344)	
Child wasted	1,851	0.020	154	0.013	0.989
		(0.140)		(0.114)	
Moderate acute malnutrition	1,851	0.016	154	0.006	0.507
		(0.124)		(0.081)	
Severe acute malnutrition	1,851	0.004	154	0.006	0.432
		(0.066)		(0.081)	
Children born in the last 24 months who were put to the breast within one hour	2,173	0.764	186	0.769	0.745
		(0.424)		(0.423)	
Infants 0-5 months of age who are fed exclusively with breast milk	1,129	0.013	97	0.031	0.599
		(0.115)		(0.174)	
Children 12-15 months of age who are fed breast milk	102	0.833	<20	0.778	0.926
		(0.375)		(0.441)	
Infants 6-8 months of age who receive solid, semi-solid or soft foods	461	0.696	37	0.703	0.989
		(0.460)		(0.463)	
Children 6-23 months of age who consume 4 or more food groups	1,038	0.214	89	0.292	0.120
		(0.410)		(0.457)	
Number of food groups (of 7) children 6-23 months of age consume	1,038	2.366	89	2.438	0.843
		(1.406)		(1.522)	
Children 6-23 months of age who meet the minimum meal frequency	1,044	0.455	89	0.427	0.596
		(0.498)		(0.497)	

	N	All	N	Eligible for Household Randomization	P-value
Women's Dietary Diversity Score (1-9)	2,833	3.529	276	3.696	0.055
		(1.112)		(1.059)	

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. To be in the sub-set households had to reside in a treatment village and the primary female and the primary male had to each own a distinct mobile phone. Standard deviations are in parentheses.

For the 18 primary outcomes and child anthropometry measures displayed in , we never reject the null hypothesis of no difference between the overall sample mean and the mean for households eligible for the sub-randomization at the 5% level. Though the lack of statistical power certainly contributes to the lack of differences, there is also no clear directional pattern between the means for the two groups. Though by no means definitive, the lack of differences uncovered in offer some evidence that the treatment effects we estimate to answer research question 5 may be relevant for the broader sample.

7.3.3 Secondary Outcomes

Table 7.3 presents the same statistics for the IYCF knowledge and beliefs outcomes. There is no significant difference in IYCF knowledge for primary females but there is some evidence that primary males in households eligible for the sub-randomization have lower levels of knowledge about IYCF practices. On average, primary males in households eligible for the sub-randomization answered 0.38 additional IYCF knowledge related questions correctly.

Table 7.3: Secondary Outcomes, by mNutrition household randomization eligibility status

	N	All	N	Eligible for Household Randomization	P-value
Infant and Young Child Feeding Knowledge (female)					
Percentage of correct answers	2,833	55.948	276	57.477	0.363
		(17.749)		(17.182)	
Infant and Young Child Feeding Knowledge (male)					
Percentage of correct answers	1,508	48.463	276	45.652	0.041
		(24.165)		(20.555)	

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. To be in the sub-set households had to reside in a treatment village and the primary female and the primary male had to each own a distinct mobile phone. Standard deviations are in parentheses.

7.4 Household Randomization: Testing for Balance Across Sub-Treatment Arms

The remainder of Annex F tests for balance in observable characteristics across the two sub-treatment arms (T-F and T-F+M) within the sub-sample of households that were eligible for the household randomization. As mentioned above, we assess balance both by calculating the normalized difference and by testing for a difference in means for each characteristic.

7.4.1 Household Demographics, Physical Structure, Amenities, and Wealth

By design, half of the 276 eligible households were allocated to the female only sub-treatment arm and the other half were assigned to the male and female sub-treatment arm. Of the 28 characteristics tested, we are able to reject the null hypothesis of no difference between the means for each sub-treatment group just once: for the likelihood that the primary female is a second order wife, after limiting the sample to the 23

eligible polygamous households. This is also the indicator with the largest normalized difference: -0.87, which is more than three times the 0.25 threshold in absolute value. The other wife order characteristics appear similarly imbalanced across the two groups, with women in the female only sub-treatment arm more likely to be second or fourth order wives and less likely to be first or third order wives. The normalized differences for three of the four wife order indicators are above the 0.25 threshold. However, wife order indicators are only non-missing for 23 polygamous households in the sample, and there is no obvious pattern to the differences. Pooling together all primary females who are in polygamous marriages and testing for differences in the average wife order yields a normalized difference of -0.11 and a p-value from the test of no difference in means between the treatment group and the control group of 0.82. Thus, the initially observed differences in wife order appear to be driven largely by problems related to the extremely small sample of eligible primary females, and not by true systematic differences across the two groups.

Outside of the wife order indicators, there are no normalized differences above 0.25 and we never fail to reject the null of no difference in means between the two sub-treatment groups at the 5% level. This suggests that, overall, the two treatment groups are quite well balanced with respect to demographic characteristics.

Table 7.4: Demographics, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Household size	138	5.188	138	5.572	0.167	0.172
		(2.252)		(2.339)		
Female-Headed Household	138	0.029	138	0.029	0.000	1.000
		(0.168)		(0.168)		
Age of Household Head	138	37.014	138	38.014	0.085	0.524
		(11.504)		(12.038)		
Household Head has some education	138	0.986	138	0.993	0.070	0.559
		(0.120)		(0.085)		
Household Head: Never Married	138	0.000	138	0.007	0.120	0.311
		(0.000)		(0.085)		
Household Head: Married, Monogamous	138	0.862	138	0.826	-0.100	0.384
		(0.346)		(0.380)		
Household Head: Married, Polygamous	138	0.116	138	0.159	0.126	0.231
		(0.321)		(0.367)		
Household Head's main activity is crop production	138	0.725	138	0.746	0.049	0.633
		(0.448)		(0.437)		
Age of the Primary Female	138	26.746	138	27.116	0.059	0.630
		(6.253)		(6.227)		
Primary Female has some education	138	0.957	138	0.964	0.037	0.781
		(0.205)		(0.188)		
Primary Female: Never Married	138	0.181	138	0.152	-0.078	0.590
		(0.387)		(0.360)		
Primary Female: Married, Monogamous	138	0.739	138	0.754	0.033	0.790
		(0.441)		(0.432)		
Primary Female: Married, Polygamous	138	0.080	138	0.080	0.000	1.000
		(0.272)		(0.272)		
Order of wife – first	<20	0.091	<20	0.333	0.594	0.208
		(0.302)		(0.492)		
Order of wife – second	<20	0.818	<20	0.417	-0.867	0.028
		(0.405)		(0.515)		
Order of wife – third	<20	0.000	<20	0.167	0.606	0.158

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
		(0.000)		(0.389)		
Order of wife – fourth	<20	0.091	<20	0.083	-0.026	0.951
		(0.302)		(0.289)		
Duration of primary female living with most recent husband/partner (months)	112	86.313	112	89.268	0.040	0.767
		(73.502)		(74.039)		
Primary Female's main activity is crop production	138	0.761	138	0.696	-0.146	0.212
		(0.428)		(0.462)		
Female focus child (under 12 months)	85	0.518	77	0.442	-0.152	0.365
		(0.503)		(0.500)		
Percent of school-aged children going to school	128	46.345	132	54.883	0.212	0.099
		(40.954)		(39.491)		
Total PPI score	138	44.130	138	40.978	-0.213	0.088
		(14.168)		(15.369)		
Consumer Durable Asset Index	138	0.548	138	0.896	0.155	0.221
		(2.189)		(2.283)		
Household Production Asset Index	138	0.300	138	0.635	0.182	0.097
		(1.554)		(2.082)		
Household Livestock Asset Index	138	-0.000	138	0.078	0.083	0.460
		(0.710)		(1.118)		
Household Total Asset Index	138	0.630	138	1.079	0.177	0.139
		(2.450)		(2.626)		
Household has access to improved sanitation	138	0.797	138	0.790	-0.018	0.878
		(0.404)		(0.409)		
Household has access to improved drinking water sources	138	0.790	138	0.804	0.036	0.775
		(0.409)		(0.398)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

7.4.2 Primary Outcomes and Child Anthropometry

Table 7.5 continues by displaying the balance measures for the primary outcomes and child anthropometry. If anything, children appear to be taller and heavier, on average, in the female only sub-treatment arm, though these differences are not even marginally statistically significant and the normalized differences are below 0.25. In fact, the only variable with a normalized difference greater than 0.25 is the likelihood that children are stunted, with children in the male and female sub-treatment arm more likely to be stunted than those in the female only arm (32.9% compared to 21.0%).

For some of the IYCF indicators that are missing for all children outside of a narrow age range, the sample size drops below 50 total observations across both groups, making it difficult to interpret the difference in means tests with much confidence. That said, the normalized differences which are less sensitive to changes in sample size, do not suggest that there is important imbalance in the primary outcomes or child anthropometry measures.

Table 7.5: Primary Outcomes, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Child age in months (under 12 months)	85	5.280	77	5.323	0.012	0.936
		(3.498)		(3.399)		

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Male child (under 12 months)	85	0.482	77	0.558	0.152	0.365
		(0.503)		(0.500)		
Child height	81	61.805	73	61.900	0.015	0.923
		(5.925)		(6.753)		
Child weight	81	6.558	73	6.634	0.046	0.768
		(1.573)		(1.737)		
Child height-for-age Z-score	81	-1.260	73	-1.400	-0.108	0.491
		(1.017)		(1.516)		
Child weight-for-age Z-score	81	-0.581	73	-0.665	-0.069	0.627
		(1.156)		(1.266)		
Child weight-for-height Z-score	81	0.390	73	0.489	0.086	0.564
		(1.206)		(1.065)		
Child stunted	81	0.210	73	0.329	0.269	0.135
		(0.410)		(0.473)		
Child underweight	81	0.148	73	0.123	-0.072	0.611
		(0.357)		(0.331)		
Child wasted	81	0.025	73	0.000	-0.224	0.154
		(0.156)		(0.000)		
Moderate acute malnutrition	81	0.012	73	0.000	-0.157	0.319
		(0.111)		(0.000)		
Severe acute malnutrition	81	0.012	73	0.000	-0.157	0.319
		(0.111)		(0.000)		
Children born in the last 24 months who were put to the breast within one hour	95	0.779	91	0.758	-0.049	0.740
		(0.417)		(0.431)		
Infants 0-5 months of age who are fed exclusively with breast milk	49	0.020	48	0.042	0.122	0.639
		(0.143)		(0.202)		
Children 12-15 months of age who are fed breast milk	<20	0.800	<20	0.750	-0.105	0.881
		(0.447)		(0.500)		
Infants 6-8 months of age who receive solid, semi-solid or soft foods	<20	0.706	20	0.700	-0.013	0.969
		(0.470)		(0.470)		
Children 6-23 months of age who consume 4 or more food groups	46	0.283	43	0.302	0.043	0.845
		(0.455)		(0.465)		
Number of food groups (of 7) children 6-23 months of age consume	46	2.391	43	2.488	0.064	0.770
		(1.598)		(1.454)		
Children 6-23 months of age who meet the minimum meal frequency	46	0.435	43	0.419	-0.032	0.890
		(0.501)		(0.499)		
Women's Dietary Diversity Score (1-9)	138	3.638	138	3.754	0.109	0.413
		(1.080)		(1.038)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

7.4.3 Secondary Outcomes

Table 7.6 displays the measures of balance for the IYCF knowledge and beliefs secondary outcome. While both primary females and primary males in the male and female sub-treatment arm perform slightly worse on the IYCF knowledge questions, the differences are small in magnitude. The normalized difference for the primary female IYCF knowledge is 0.03 and the p-value from the t-test of no difference between the

means in the female only group and the female and male group is 0.83. Similarly, the normalized difference for primary male IYCF knowledge is just -0.17 with a p-value of 0.16.

Table 7.6: Secondary Outcomes, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Infant and Young Child Feeding Knowledge (female)						
Percentage of correct answers	138	57.708	138	57.246	-0.027	0.829
		(16.242)		(18.129)		
Infant and Young Child Feeding Knowledge (male)						
Percentage of correct answers	138	47.431	138	43.874	-0.173	0.164
		(19.237)		(21.717)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 7.7: Sources of information and trust likelihood on health and nutrition, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
The following is a source of information:						
Spouse	114	0.623	115	0.670	0.097	0.482
		(0.487)		(0.472)		
Family	138	0.478	138	0.478	0.000	1.000
		(0.501)		(0.501)		
Friends/neighbours	138	0.500	138	0.572	0.145	0.189
		(0.502)		(0.497)		
Automated text messages	138	0.355	138	0.283	-0.155	0.197
		(0.480)		(0.452)		
Government health workers	138	0.978	138	0.964	-0.086	0.390
		(0.146)		(0.188)		
Non-government health facilities	138	0.493	138	0.420	-0.145	0.232
		(0.502)		(0.495)		
TV/Radio/Posters	138	0.761	138	0.732	-0.066	0.567
		(0.428)		(0.445)		
Traditional health workers	138	0.087	138	0.094	0.025	0.840
		(0.283)		(0.293)		
Non-government health workers	138	0.384	138	0.471	0.176	0.188
		(0.488)		(0.501)		
Community health workers	138	0.572	138	0.630	0.118	0.311
		(0.497)		(0.484)		
*The following is the **primary source** of information:*						
Spouse	138	0.181	138	0.225	0.108	0.341
		(0.387)		(0.419)		
Family	138	0.080	138	0.094	0.051	0.701
		(0.272)		(0.293)		
Friends/neighbors	138	0.043	138	0.007	-0.231	0.048
		(0.205)		(0.085)		
Automated text messages	138	0.022	138	0.029	0.046	0.576

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
		(0.146)		(0.168)		
Government health workers	138	0.841	138	0.775	-0.166	0.172
		(0.367)		(0.419)		
Non-government health facilities	138	0.123	138	0.101	-0.069	0.574
		(0.330)		(0.303)		
TV/Radio/Posters	138	0.072	138	0.109	0.126	0.330
		(0.260)		(0.312)		
Traditional health workers	138	0.000	138	0.007	0.120	0.311
		(0.000)		(0.085)		
Non-government health workers	138	0.036	138	0.051	0.071	0.583
		(0.188)		(0.220)		
Community health workers	138	0.087	138	0.094	0.025	0.823
		(0.283)		(0.293)		
*Agree they can **trust** the following source of information:*						
Spouse	114	0.851	115	0.852	0.004	0.980
		(0.358)		(0.356)		
Family	138	0.696	138	0.681	-0.031	0.793
		(0.462)		(0.468)		
Friends/neighbors	138	0.609	138	0.609	0.000	1.000
		(0.490)		(0.490)		
Automated text messages	138	0.870	138	0.812	-0.158	0.177
		(0.338)		(0.392)		
Government health workers	138	0.993	138	0.964	-0.199	0.096
		(0.085)		(0.188)		
Private clinic/hospital	138	0.877	138	0.870	-0.022	0.863
		(0.330)		(0.338)		
TV/radio/posters	138	0.899	138	0.841	-0.172	0.209
		(0.303)		(0.367)		
Traditional health worker	138	0.065	138	0.094	0.107	0.347
		(0.248)		(0.293)		
Non-government health worker	138	0.732	138	0.804	0.172	0.134
		(0.445)		(0.398)		
Community health worker	138	0.899	138	0.942	0.160	0.269
		(0.303)		(0.235)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 7.8: Mobile phone access and usage (female), by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Owns a mobile phone	138	1.000	138	1.000		
		(0.000)		(0.000)		
Used a mobile phone in the last 14 days to						
make calls	138	0.906	138	0.855	-0.156	0.157
		(0.293)		(0.353)		
receive calls	138	0.964	138	0.971	0.041	0.736

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
		(0.188)		(0.168)		
write text messages	138	0.674	138	0.703	0.062	0.610
		(0.470)		(0.459)		
receive text messages	138	0.884	138	0.891	0.023	0.860
		(0.321)		(0.312)		
send mobile money	138	0.130	138	0.145	0.042	0.730
		(0.338)		(0.353)		
receive mobile money	138	0.181	138	0.196	0.037	0.777
		(0.387)		(0.398)		
use mobile internet	138	0.058	138	0.036	-0.102	0.384
		(0.235)		(0.188)		
No mobile phone use in the last 14 days	138	0.007	138	0.000	-0.120	0.325
		(0.085)		(0.000)		
Mobile phone used often in the last 14 days	137	0.869	138	0.877	0.025	0.846
		(0.339)		(0.330)		
Used mobile phone to receive health advice	138	0.130	138	0.152	0.062	0.581
		(0.338)		(0.360)		
Received automatic text messages (nutrition information) in past 2 years	138	0.072	138	0.123	0.171	0.183
		(0.260)		(0.330)		
Amount spent on airtime on all phones in an average month (Tz shillings)	138	3,505.797	138	3,551.449	0.011	0.935
		(3,403.793)		(4,707.072)		
Charges phone at home	138	0.565	138	0.580	0.029	0.804
		(0.498)		(0.495)		
Takes less than 30 minutes to get to the nearest place to charge their phone	60	0.983	58	0.983	-0.004	0.981
		(0.129)		(0.131)		
Amount spent on charging phone in an average month (Tz shillings)	60	1,185.000	58	1,675.862	0.299	0.137
		(1,332.746)		(1,898.227)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 7.9: Mobile phone access and usage (male), by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Owns a mobile phone	138	1.000	138	1.000		
		(0.000)		(0.000)		
Used a mobile phone in the last 14 days to						
make calls	138	0.993	138	0.978	-0.121	0.323
		(0.085)		(0.146)		
receive calls	138	1.000	138	0.971	-0.243	0.098
		(0.000)		(0.168)		
write text messages	138	0.783	138	0.754	-0.068	0.556
		(0.414)		(0.432)		
receive text messages	138	0.957	138	0.928	-0.124	0.275
		(0.205)		(0.260)		
send mobile money	138	0.275	138	0.333	0.126	0.342
		(0.448)		(0.473)		
receive mobile money	138	0.290	138	0.399	0.229	0.045

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
		(0.455)		(0.491)		
use mobile internet	138	0.101	138	0.116	0.046	0.689
		(0.303)		(0.321)		
No mobile phone use in the last 14 days	138	0.014	138	0.000	-0.171	0.157
		(0.120)		(0.000)		
Mobile phone used often in the last 14 days	136	0.963	138	0.942	-0.100	0.411
		(0.189)		(0.235)		
Used mobile phone to receive health advice	138	0.101	138	0.080	-0.076	0.536
		(0.303)		(0.272)		
Received automatic text messages (nutrition information) in past 2 years	138	0.116	138	0.036	-0.303	0.007
		(0.321)		(0.188)		
Amount spent on airtime on all phones in an average month (Tz shillings)	138	9,852.891	138	9,742.754	-0.011	0.936
		(9,520.084)		(10,120.665)		
Charges phone at home	138	0.580	138	0.580	0.000	1.000
		(0.495)		(0.495)		
Takes less than 30 minutes to get to the nearest place to charge their phone	58	0.983	58	0.966	-0.108	0.551
		(0.131)		(0.184)		
Amount spent on charging phone in an average month (Tz shillings)	58	2,006.897	58	1,875.862	-0.062	0.709
		(1,964.765)		(2,270.197)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 7.10: Child Health, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
General health rated good, very good, or excellent	179	0.849	182	0.857	0.022	0.853
		(0.359)		(0.351)		
Was ill with a fever in the past 14 days	179	0.251	182	0.170	-0.199	0.138
		(0.435)		(0.377)		
Was ill with a cough in the past 14 days	179	0.369	182	0.308	-0.129	0.288
		(0.484)		(0.463)		
Was ill with diarrhea in the past 14 days	179	0.190	182	0.099	-0.260	0.024
		(0.393)		(0.299)		
Duration of diarrhea in the past 14 days (if child had diarrhea)	34	4.529	<20	4.333	-0.058	0.839
		(2.852)		(3.789)		
Received BCG vaccine	178	0.978	179	0.978	0.001	0.995
		(0.149)		(0.148)		
Received polio vaccine	178	0.961	180	0.978	0.099	0.383
		(0.195)		(0.148)		
Received DPT-HB vaccine	177	0.915	180	0.939	0.091	0.437
		(0.279)		(0.240)		
Received MMR or measles injection (9 months or older)	115	0.913	120	0.933	0.076	0.519
		(0.283)		(0.250)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample of children born in 2011 or later who are in the household. Standard deviations are in parentheses.

Table 7.11: Maternal Health, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Received antenatal care	123	1.000	123	1.000		
		(0.000)		(0.000)		
Received antenatal care at home	123	0.000	123	0.000		
		(0.000)		(0.000)		
Received antenatal care at a government health facility	123	0.959	123	0.951	-0.039	0.805
		(0.198)		(0.216)		
Received antenatal care at a private hospital/clinic	123	0.024	123	0.008	-0.128	0.312
		(0.155)		(0.090)		
Received antenatal care at a religious or voluntary hospital	123	0.016	123	0.041	0.146	0.403
		(0.127)		(0.198)		
Number of months pregnant at time of first ANC visit	123	3.537	123	3.463	-0.063	0.692
		(1.140)		(1.169)		
Number of ANC visits	123	4.276	123	4.301	0.018	0.908
		(1.301)		(1.425)		
Blood pressure measured during antenatal care	122	0.877	123	0.837	-0.113	0.470
		(0.330)		(0.371)		
Urine sample collected during antenatal care	122	0.762	123	0.683	-0.177	0.247
		(0.427)		(0.467)		
Blood sample collected during antenatal care	123	0.967	123	0.951	-0.082	0.504
		(0.178)		(0.216)		
Told about signs of pregnancy complications	122	0.697	120	0.650	-0.099	0.509
		(0.462)		(0.479)		
Received tetanus injections	123	0.862	122	0.877	0.045	0.739
		(0.347)		(0.330)		
Received iron supplements	123	0.919	123	0.911	-0.029	0.836
		(0.274)		(0.287)		
Had difficulty with vision during the daylight	123	0.065	123	0.008	-0.305	0.017
		(0.248)		(0.090)		
Suffered from night blindness	123	0.033	123	0.016	-0.105	0.420
		(0.178)		(0.127)		
Slept under a bed net	123	0.862	123	0.878	0.048	0.688
		(0.347)		(0.329)		
Took anti-malaria medication	123	0.870	123	0.943	0.252	0.083
		(0.338)		(0.233)		
Instructed to attend a follow up visit at nearest dispensary (after child's birth)	123	0.837	123	0.846	0.022	0.874
		(0.371)		(0.363)		
Place of delivery: Home	122	0.016	123	0.033	0.104	0.567
		(0.128)		(0.178)		
Place of delivery: Government health facility	122	0.918	123	0.911	-0.027	0.868
		(0.275)		(0.287)		
Place of delivery: Private hospital/clinic	122	0.041	123	0.016	-0.148	0.250
		(0.199)		(0.127)		
Place of delivery: Religious or voluntary hospital	122	0.025	123	0.041	0.090	0.597
		(0.156)		(0.198)		
Had a postnatal follow up within the first week of the child's birth	123	0.854	123	0.878	0.071	0.623
		(0.355)		(0.329)		

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Number of postnatal follow ups within 28 days of the child's birth	123	1.561	123	1.732	0.192	0.135
		(0.888)		(0.888)		
Number of postnatal follow ups within 42 days of the child's birth	123	2.374	123	2.748	0.191	0.160
		(1.169)		(2.511)		
Visited at home by community health worker during the first few days after child	123	0.195	123	0.236	0.099	0.544
		(0.398)		(0.426)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample of all births in 2013 or later. Standard deviations are in parentheses.

Table 7.12: HIV/AIDS Awareness, by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Heard of illness called HIV/AIDS	138	0.978	138	1.000	0.210	0.082
		(0.146)		(0.000)		
Tested for the HIV virus	138	1.000	138	0.993	-0.120	0.319
		(0.000)		(0.085)		
Received results for the HIV virus test (of those tested)	138	1.000	137	1.000		
		(0.000)		(0.000)		
Partner tested for the HIV virus	111	0.955	115	0.957	0.008	0.958
		(0.208)		(0.205)		
Heard of antiretroviral therapy	138	0.797	138	0.891	0.261	0.024
		(0.404)		(0.312)		
Knows where to receive antiretroviral therapy	110	0.745	123	0.764	0.043	0.754
		(0.438)		(0.426)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 7.13: Desired fertility (female), by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Primary female is currently pregnant	138	0.370	138	0.420	0.104	0.394
		(0.484)		(0.495)		
Gestation of pregnancy (weeks)	51	25.804	58	26.552	0.103	0.609
		(7.125)		(7.368)		
Time since respondent found out about their pregnancy (weeks)	51	16.706	58	16.741	0.004	0.985
		(8.707)		(9.229)		
Time respondent would like to wait before conceiving their next child (years)	88	4.489	87	4.716	0.129	0.374
		(1.661)		(1.854)		
Would like to have another child	138	0.638	138	0.616	-0.045	0.737
		(0.482)		(0.488)		
Would not like to have another child	138	0.341	138	0.348	0.015	0.912
		(0.476)		(0.478)		
Is undecided about having another child	138	0.022	138	0.036	0.086	0.484
		(0.146)		(0.188)		
Will use a contraceptive method after their pregnancy	48	0.875	55	0.836	-0.109	0.587
		(0.334)		(0.373)		

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Is using a contraceptive method	87	0.483	80	0.463	-0.040	0.792
		(0.503)		(0.502)		
Method of contraceptive being used: IUD	42	0.190	37	0.054	-0.420	0.061
		(0.397)		(0.229)		
Method of contraceptive being used: Injectables	42	0.643	37	0.730	0.186	0.445
		(0.485)		(0.450)		
Method of contraceptive being used: Pill	42	0.119	37	0.081	-0.125	0.581
		(0.328)		(0.277)		
Method of contraceptive being used: Condom	42	0.095	37	0.027	-0.284	0.216
		(0.297)		(0.164)		
Method of contraceptive being used: Female Condom	42	0.000	37	0.000		
		(0.000)		(0.000)		
Method of contraceptive being used: Sterilization	42	0.000	37	0.027	0.232	0.326
		(0.000)		(0.164)		
Number of living children	138	2.217	138	2.326	0.060	0.647
		(1.799)		(1.801)		
Number of additional children wanted	91	2.440	90	2.444	0.004	0.978
		(1.231)		(1.299)		
Number of additional male children wanted	91	0.846	90	0.700	-0.160	0.352
		(0.965)		(0.854)		
Number of additional female children wanted	91	0.769	90	0.678	-0.104	0.506
		(0.944)		(0.819)		
Number of additional children wanted where the sex would not matter	91	0.824	90	1.067	0.180	0.310
		(1.111)		(1.549)		
Partner would like to have the same number of children	114	0.544	115	0.478	-0.131	0.267
		(0.500)		(0.502)		
Partner would like to have more children	114	0.035	115	0.061	0.120	0.366
		(0.185)		(0.240)		
Partner would like to have fewer children	114	0.018	115	0.035	0.108	0.435
		(0.132)		(0.184)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Table 7.14: Desired fertility (male), by mNutrition sub-treatment status, Eligible Sample

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
Wife/partner is currently pregnant	111	0.378	108	0.389	0.022	0.882
		(0.487)		(0.490)		
Gestation of pregnancy (weeks)	41	26.732	42	26.262	-0.066	0.767
		(6.896)		(7.408)		
Time since respondent found out about their pregnancy (weeks)	42	16.476	42	15.405	-0.122	0.595
		(9.402)		(8.157)		
Time respondent would like to wait before conceiving their next child (years)	69	3.790	78	4.067	0.180	0.252
		(1.441)		(1.638)		
Would like to have another child	111	0.604	108	0.630	0.053	0.727
		(0.491)		(0.485)		
Would not like to have another child	111	0.360	108	0.269	-0.198	0.182

	N	Female only sub-treatment	N	Male and female sub-treatment	Normalized Difference	P-value
		(0.482)		(0.445)		
Is undecided about having another child	111	0.036	108	0.102	0.261	0.034
		(0.187)		(0.304)		
Will use a contraceptive method after their pregnancy	40	0.750	39	0.744	-0.015	0.956
		(0.439)		(0.442)		
Is using a contraceptive method	69	0.290	66	0.273	-0.038	0.844
		(0.457)		(0.449)		
Method of contraceptive being used: IUD	20	0.000	<20	0.000		
		(0.000)		(0.000)		
Method of contraceptive being used: Injectables	20	0.300	<20	0.444	0.294	0.337
		(0.470)		(0.511)		
Method of contraceptive being used: Pill	20	0.000	<20	0.000		
		(0.000)		(0.000)		
Method of contraceptive being used: Condom	20	0.700	<20	0.500	-0.406	0.186
		(0.470)		(0.514)		
Method of contraceptive being used: Female Condom	20	0.050	<20	0.000	-0.316	0.329
		(0.224)		(0.000)		
Method of contraceptive being used: Sterilization	20	0.000	<20	0.000		
		(0.000)		(0.000)		
Number of living children	111	3.387	108	3.148	-0.070	0.627
		(3.980)		(2.696)		
Number of additional children wanted	71	2.408	79	2.354	-0.032	0.830
		(1.489)		(1.868)		
Number of additional male children wanted	71	0.972	79	0.861	-0.107	0.590
		(1.028)		(1.047)		
Number of additional female children wanted	71	0.620	79	0.481	-0.185	0.300
		(0.781)		(0.714)		
Number of additional children wanted where the sex would not matter	71	0.817	79	1.013	0.109	0.543
		(1.570)		(2.003)		
Partner would like to have the same number of children	111	0.514	108	0.472	-0.082	0.557
		(0.502)		(0.502)		
Partner would like to have more children	111	0.054	108	0.009	-0.257	0.055
		(0.227)		(0.096)		
Partner would like to have fewer children	111	0.000	108	0.019	0.193	0.163
		(0.000)		(0.135)		

Notes: Estimates from the mNutrition Tanzania Baseline Survey sample. Standard deviations are in parentheses.

Annex G Terms of Reference

Call-down Contract

Terms of Reference

PO 6420: External evaluation of mobile phone technology based nutrition and agriculture advisory services in Africa and South Asia

Introduction

DFID (Research and Evidence Division) wishes to commission an external impact evaluation of mNutrition, a mobile phone technology based nutrition and agricultural advisory service for Africa and South Asia. mNutrition is a programme supported by DFID that, through business and science partnerships, aims to build sustainable business models for the delivery of mobile phone technology based advisory services that are effective in improving nutrition and agricultural outcomes.
mNutrition is primarily designed to use mobile phone based technologies to increase the access of rural communities to nutrition and agriculture related information. The initiative aims to improve knowledge among rural farming communities especially women and support beneficial behaviour change as well as increasing demand for nutrition and agriculture extension services. The mNutrition initiative launched in September 2013 will work in 10 countries in Africa (Cote d'Ivoire, Ghana, Malawi, Mozambique, Nigeria, Tanzania, Kenya, Rwanda, Uganda, Zambia) and four countries in South Asia (Bangladesh, India, Pakistan and Sri Lanka). The desired impact of mNutrition will be improved nutrition, food security and livelihoods of the poor.

Mobile phone based services have been endorsed by WHO as an effective strategy for behaviour change and for driving adherence to anti-retroviral treatment protocols (Horvath, Azman, Kennedy and Rutherford 2012). There is currently scant evidence on the impact and cost-effectiveness of mobile phone technology based services for nutrition and agriculture and on the sustainability of different business models for their provision. A rigorous evaluation of mobile phone technology based nutrition services would add significantly to the current evidence base. An external evaluation team managed by the Evaluator, independent of the programme delivery mechanism, will conduct an assessment of the impact, cost-effectiveness and sustainability of mobile phone technology based information and behaviour change messages for nutrition and agriculture.

Background to mNutrition

Introduction

Undernutrition is a major challenge to human and economic development globally. It is estimated that almost one billion people face hunger and are unable to get enough food to meet their dietary needs. Agriculture is a major source of livelihood in many poor countries and the sector has a potentially critical role in enhancing health, specifically maternal and child health and nutritional status. A well-developed agriculture sector will deliver increased and diversified farm outputs (crops, livestock, non-food products) and this may enhance food and nutrition security directly through increased access to and consumption of diverse food, or indirectly through greater profits to farmers and national wealth. Better nutrition and health of farmers fosters their agricultural and economic productivity. Current agricultural and health systems and policies are not meeting current and projected future global food, nutrition and health needs.

Despite major investment in agricultural and nutrition research and its uptake and application, there is significant social and geographic inequality in who benefits from these investments. Furthermore, in many developing countries, public extension systems for agriculture, health and nutrition are inefficient, have limited capacity and have a poor track record of delivery, especially in terms of supporting women and girls and the most marginalised populations (Alston, Wyatt, Pardey, Marra and Chan-Kang 2000; Anderson 2007; IFPRI 2010; Van den Berg and Jiggins 2007).

Several research and mobile network operators (MNOs) are testing a range of information and communication technology (ICT) solutions for improving access to a wide range of information and advisory

services. Mobile phone based technologies are among the most promising ICT strategies, although current initiatives in nutrition are relatively small and fragmented.

What is mNutrition?

Enhancing access to the results of nutrition and agricultural research and development is potentially critical for improving the nutrition, health and livelihoods of smallholders and rural communities. mNutrition will harness the power of mobile phone based technologies and the private sector to improve access to information on nutrition, health and agricultural practices especially for women and farmers (both male and female). Specifically, mNutrition will initiate new partnerships with business and science to deliver a range of services including:

- An open-access database of nutrition and agriculture messages for use in mobile phone based communication (for example, information and behaviour change messages on practices and interventions that are known to have a direct impact on nutrition or an indirect impact via for example agriculture);
- A suite of mobile phone based nutrition and agriculture information, extension and registration services designed to: improve knowledge and generate beneficial behaviour change in nutrition and agriculture; increase demand for nutrition, health and agriculture goods and services; register and identify target populations for support; and, using real-time monitoring, support the conduct of nutrition risk assessments by community health workers.

The impacts of mNutrition are expected to include improved nutrition, food security and livelihoods of the poor, especially women in 10 countries in Africa (Cote d'Ivoire, Ghana, Kenya, Malawi, Mozambique, Nigeria, Rwanda, Tanzania, Uganda and Zambia) and 4 countries in South Asia (Bangladesh, India, Pakistan and Sri Lanka). This impact will result from the increased scale and sustainability of mobile phone based nutrition and agricultural-based information services, delivered through robust public private partnerships in each country.

mNutrition has two major outcomes. One outcome will be cost-effective, sustainable business models for mobile phone enabled nutrition and agriculture services to 3 million households in 10 countries in Africa and 4 countries in South Asia that can be replicated in other countries. Linked to this outcome, the second outcome will expect these services to result in new knowledge, behaviour change and adoption of new practices in the area of agriculture and nutrition practices among the users of these mobile phone based services.

These outcomes will be achieved through four outputs:

- Improved access to relevant mobile based health, nutrition and agricultural advisory services for 3 million poor people and community health workers across 10 SSA and 4 Asian countries;
- Launch and scaling of mobile phone based health, nutrition and agricultural advisory services targeted to poor people and community health workers;
- Generation and dissemination of high quality research and evidence on the impact, cost-effectiveness and sustainability of mobile phone based advisory services in nutrition and agriculture in South Asia and SSA; and
- Development of locally relevant content for mobile phone technology based agriculture and nutrition services meeting demands from users and community health workers.

In terms of promoting behaviour change and/or adoption of new practices, mNutrition will seek to achieve changes in one or more of the following areas:

- Adoption of new agricultural practices that are nutrition sensitive, improve agricultural productivity and utilise post-harvest technologies
- Changes in nutrition practices in either one or several knowledge domains including improved maternal nutrition practices during pregnancies; infant and young child feeding practice; and micro-nutrient supplementation to children at risk (i.e. Vitamin A, Zinc and Oral Rehydration Solution (ORS)).

mNutrition has started implementation from September 2013. For the 2 countries selected for the impact evaluation (Tanzania and Ghana), mobile network operators and content providers have been identified through a competitive process during the first half of 2014. The MNOs and content providers started developing and launching their services during the 4th quarter of 2014 and early 2015. The mobile phone based advisory services are expected to run at least till 3rd quarter of 2018.

mNutrition Project Coordination

DFID support to mNutrition will be channelled to GSMA, as well as directly to this associated independent external impact evaluation. GSMA is a global body that represents the interests of over 800 mobile operators. GSMA already works with the major mobile operators across Africa, (including Airtel, MTN, SafariCom/VodaCom) with a collective mobile footprint of more than 67% of total African connections. GSMA has a number of existing development initiatives, including mHealth and mFarmer, that are part of GSMA's Mobile for Development which brings together mobile operator members, the wider mobile industry and the development community to drive commercial mobile services for underserved people in emerging markets. GSMA will provide technical assistance to mobile phone operators, and support new partnerships with content providers to develop and scale up new nutrition and agriculture message services. GSMA will ensure sharing of best practices and promote wider replication and uptake of effective business models.

Objective and Main Questions

The objective of this work is to conduct an external evaluation of the impacts and cost-effectiveness of the nutrition and agriculture advisory services provided by mNutrition compared to alternative advisory services available in the two selected countries (Ghana and Tanzania), with particular attention paid to gender and poverty issues. The impact assessment is required to answer the following questions that relate to impact, cost-effectiveness and commercial viability:

- What are the impacts and cost-effectiveness of mobile phone based nutrition and agriculture services on nutrition, health and livelihood outcomes, especially among women, children and the extreme poor?
- How effective are mobile phone based services in reaching, increasing the knowledge, and changing the behaviour, of the specific target groups?
- Has the process of adapting globally agreed messages to local contexts led to content which is relevant to the needs of children, women and poor farmers in their specific context?
- What factors make mobile phone based services effective in promoting and achieving behaviour change (if observed) leading to improved nutrition and livelihood outcomes?
- How commercially viable are the different business models being employed at country level?
- What lessons can be learned about best practices in the design and implementation of mobile phone based nutrition services to ensure a) behaviour change and b) continued private sector engagement in different countries?

Further evaluation questions related to other aims of mNutrition will be addressed in at least 1 country (either Ghana and/or Tanzania):

- Are mobile phone based services a cost-effective way to register and identify at risk populations to target with nutrition support?
- Are mobile phone based services a cost-effective way for community health workers to improve the quality and timeliness of data surveillance (a core set of nutrition-related indicators)?

The content for the mobile phone based advisory services will be based on international best practices and widely endorsed protocols (i.e. by the World Health Organisation) and evidence-based nutrition-sensitive agricultural practices identified by international experts. Through an iterative multi-stakeholder process, international and country experts will localise and adapt the content to make it relevant to the specific target audience in the 14 countries. The adapted content and nature of messages is expected to vary across specific target audiences within and across countries. The main purpose of assessing the relevance of the content is not to evaluate the overall health and nutrition content but on how this content has been localised and adapted and to what extent the needs of the specific target groups within their particular context have been met.

In assessing the commercial viability, it is recognised that evaluating the sustainability/long-term financial viability of the mobile phone based advisory services will be difficult as mobile network operators may not be willing to provide this potentially commercially sensitive information. Therefore, GSMA will provide support through its access to aggregated confidential financial results of the mobile network operators providing the service. GSMA will provide a financial summary report on the commercial viability of the business models without compromising the commercial sensitivity of the data for the mobile network operators. The evaluator will assess and validate commercial sustainability through an analysis of the aggregated information provided by GSMA and additional qualitative business analysis approaches.

The Evaluator has the option of proposing refinements of the existing evaluation questions during the inception phase as part of developing the research protocol. These suggestions will be considered by the

Steering Committee and an independent peer review during the review of the research protocol as part of the inception phase.

Output

The output of this work will be new and robust evidence on the impact, cost-effectiveness and commercial viability of mobile phone based advisory services focusing on nutrition and agriculture delivered by public and private partners, and including the development of robust methodological approaches to impact assessment of phone based advisory services.

Recipient

The primary recipient of this work will be DFID, with the beneficiaries being GSMA, governments, international agencies, foundations, MNOs and other private companies and civil society involved in policies and programmes in nutrition and agriculture that are aimed at improving nutritional, health and agricultural outcomes. The findings of this impact evaluation are intended as global public goods.

Scope and timeline

The scope of this work is to:

- Develop a research protocol for the external evaluation of mNutrition;
- Design and undertake an external evaluation of mNutrition in two countries: Ghana and Tanzania;
- Contribute to the communication of the learning agenda, evaluation strategy and evaluation results.

The evaluation will be in two of the 14 mNutrition target countries; Ghana and Tanzania. These countries have been selected based on the phased start-up of mNutrition programme activities. The focus and approach in the two respective countries will be different allowing for a comparison of the effectiveness of approaches applied. In Tanzania, mNutrition will focus on mobile phone technology based nutrition and health services and registration and identification of target population. In Ghana, the mobile phone technology will focus on nutrition and agriculture sensitive services.

In terms of coverage in number of people being targeted for these services, in total 3 million people will be reached through mNutrition; including 2 million for nutrition sensitive agriculture advisory messages in 4 Asian and at least 2 African countries and about 1 million beneficiaries for mobile phone based nutrition services in 10 countries in SSA.

The evaluation contract period will be September 2014 to 31st December 2019. The development of the research protocol must be completed by month 4 for review and approval by DFID. Full details on tasks and deliverables are provided in sections below.

Statement on the design of the mNutrition evaluation

The evaluation design is expected to measure the impact, cost-effectiveness and commercial viability of mNutrition, using a mixed methods evaluation design and drawing on evidence from two case study countries and the M&E system of the programme. Overall, the proposed design should ensure that the evidence from the two case study countries has high internal validity and addresses the priority evidence gaps identified in the Business Case. Being able to judge the generalisability/replicability of lessons learned from the programme is of equal importance and so a credible approach to generalization and external validity will be an important component of the overall evaluation design. The final evaluation design and methodology to generate robust evidence will be discussed in detail with DFID and GSMA before implementation.

For assessing cost-effectiveness, the Evaluator will further fine-tune their proposed evaluation approach and outline their expectations in terms of data they will require from implementers. A theory based evaluation design, using mixed methods for evaluating the impact has been proposed. During the inception phase, the Evaluator will put forward a robust evaluation design for the quantitative work, either an experimental or a quasi-experimental method, with a clear outline of the strengths and limitations of the proposed method relative to alternatives. During the inception phase, the Evaluator is also expected to identify clearly what will be the implications of the design for implementers in terms of how the overall programme would be designed and implemented and for evidence to be collected in the programme's monitoring system. The Evaluator will also assess the degree to which it is realistic to assess impacts by early 2019 for a programme where implementation started mid 2015 and, if there are challenges, how these would be managed.

The Evaluator, in its 6 monthly reports, will be required to provide information to feed into the DFID Annual Review and Project Completion Report of mNutrition.

Gender and inclusiveness

The impact evaluation will pay particular attention to gender and other forms of social differentiation and poverty issues. From current experiences, it is clear that access to and use of mobile services is differentiated along a range of factors, including gender, poverty, geographic marginalisation, education and illiteracy levels. Therefore, the impact evaluation will look at and analyse differentiated access to and potential utilisation of mobile phone based services for improved nutrition and agricultural production. Based on the findings, it will identify opportunities and challenges in having an impact on women in general and more specifically the poor and the marginalised.

Tasks

The Evaluator will perform the following tasks:

A. Finalise a coherent and robust evaluation approach and methodology based on their proposal (inception phase)

- Conduct landscape analysis of existing experiences in mobile phone based services for nutrition and agriculture based on available publications and grey project documents to identify additional critical lessons and priorities for evidence gathering and programme design and implementation;
- Ensure that gender issues and poverty issues are well integrated into the impact evaluation design;
- Develop robust sampling frameworks, core set of indicators and research protocols that allow the consistent measurement and comparison of impacts across study countries, taking into account differences in business models and programmes as needed;
- Work closely with mNutrition programme team in GSMA to familiarise them with impact assessment methodology, discuss evaluation approaches, identify and agree on data provided by programme monitoring system and possible modifications to design;
- Identify risks to the evaluation meeting its objectives and how these risks will be effectively managed;
- Review existing evaluation questions and if deemed relevant propose refinement of existing questions and/or add other questions;
- Prepare a research protocol, including an updated workplan, project milestones and budget. The research protocol will be subject to an independent peer review organised by DFID; and
- Develop a communication plan.

B. Implement and analyse evaluations of impact, cost-effectiveness and commercial viability in accordance with established best practices

- Based upon the agreed evaluation framework, develop and test appropriate evaluation instruments which are likely to include data collection forms for households, community health workers, service providers including health and agricultural services, content providers and private sector stakeholders including mobile network operators. Instruments will involve both quantitative and qualitative methods;
- Register studies on appropriate open access study registries and publish protocols of studies where appropriate;
- Conduct baselines and end-lines, qualitative assessments and business model assessments in both of the two impact evaluation countries;
- Conduct and analyse the evaluations and present findings in two well-structured reports addressing the evaluation questions. The reports should follow standard reporting guidelines as defined by, for example, the Equator Network. Primary findings should be clearly presented along with a detailed analysis of the underlying reasons why the desired outcomes were/were not achieved;
- The Evaluating Organisation or Consortium may subcontract the administration of surveys and data entry, but not the supervision of those tasks, study design, or data analysis; and
- The country-specific mixed methods evaluation reports, cost effectiveness and business models studies and final evaluation report will be subject to an independent peer review organised by DFID.

C. Contribute to the communication of the learning agenda, impact evaluation strategy, and evaluation results.

- Develop a communication plan outlining the main outputs and key audiences;

- Conduct lessons learnt workshops in each of the 2 impact evaluation countries and key dissemination events; and
- Assist in communicating the results of the evaluation and contribute to the development and communication of lessons learnt about mobile phone based extension approaches in nutrition and agriculture.

Deliverables

The Evaluator will deliver the following outputs[53]**:**

During the design and study inception phase of maximum 4 months:
- A publishable landscape analysis report highlighting lessons learnt from existing initiatives on mobile phone based advisory services related to nutrition and agriculture by month 4;
- A updated work plan with project milestones and budget by end of month 1 (possibly adjusted based on the approved research protocol by month 4);
- A communication plan outlining the key outputs, audience and timeline for review and approval by month 4; and
- A full research protocol by month 4 for review and approval. The research protocol should be registered with appropriate open access study registries;

Interim reports:
- 4 biannual progress reports for the External Evaluation as a whole, and for each country evaluation, against milestones set out in the workplan;
  - Two desk reviews submitted by June 2016
  - Two Baseline quantitative reports submitted by April 2017
  - Two Baseline qualitative reports submitted by February 2017
  - Two Cost-effectiveness reports 1 submitted by March 2017
  - Two Business Model reports 1 submitted by March 2017
  - Two Mixed Methods Baseline reports completed by September 2017
  - Two Midline qualitative reports submitted by March 2018
- All survey data collected during the evaluation provided in a suitable format to DFID for public release.

At project's end:
- Two Endline quantitative reports submitted by June 2019
- Two Endline qualitative reports submitted by August 2019
- Two Cost-effectiveness report 2 submitted by July 2019
- Two Business Model report 2 submitted by July 2019
- Two Evaluation reports submitted by October 2019
- At least 1 article, based on the findings from the country evaluation reports, published in a research journal;
- A shared lesson learnt paper published and at least one presentation highlighting key lessons for similar initiatives of promoting mobile based technologies for providing extension services and the promotion of uptake of technologies by December 2019.

Research protocol and all final reports will be independently peer reviewed. This will be organised by DFID. Outputs are expected to be of sufficiently quality so that a synthesis of findings can be published in a leading peer-reviewed journal.

Coordination and reporting requirements

A mNutrition Advisory Group (AG) will be established for the programme which will a) provide technical oversight and b) maximise the effectiveness of the programme. The Advisory Group will meet on a bi-annual basis and comprises of representatives of DFID, NORAD and GSMA representatives and independent technical experts. The Evaluator will be managed by DFID on behalf of the mNutrition Advisory Group. The

[53] Exact timeframe of deliverables will be agreed on during the design phase as appropriate.

Evaluator will work closely with the mNutrition programme team in GSMA and its specific country implementing partners. The Evaluator will:

- Ensure coherence and lesson learning across all pilot impact assessments on the key evaluation questions and indicators identified.
- Incorporate a clear code of ethics; incorporate plans for open access publications and public access to data sets.

The Evaluator will work closely with the mNutrition project management team, in particular in the design of the overall evaluation framework and the evaluation plan for the specific project components and the countries selected for the evaluation. Collaboration and regular communication between Evaluator and mNutrition project management team and implementing partners in selected case study countries is crucial as the evaluation design may have implications for project implementation and vice versa. The mNutrition project management team will lend support in communication as requested by the Evaluator or the Advisory Group. The Evaluator will report directly to DFID who will manage the evaluation on behalf of the mNutrition Advisory Group. The main point of contact for technical matters is Louise Horner, Livelihoods Adviser and Hugh McGhie, Deputy Programme Manager for all other project related issues. The mNutrition Advisory Group will be the arbiter of any disputes between the evaluation function and the overall programme implementation.

At the end of each 6 months, the Evaluator will submit a brief report outlining key achievements against the agreed deliverables. Pre-agreed funding will then be released provided that deliverables have been achieved.

In addition to the 6 monthly reports outlined above, the Evaluator will provide information to feed into the DFID Annual Review of mNutrition. The 6 monthly reports will be a key source of information used to undertake the Annual Review and Project Completion Report for the programme. These reviews will be led by the Livelihoods Adviser and Deputy Programme Manager, in consultation with the mNutrition AG. All reviews will be made available publicly in line with HMG Transparency and Accountability Requirements.

Mandatory financial reports include an annual forecast of expenditure (the budget) disaggregated monthly in accordance with DFID's financial year April to March. This should be updated at least every quarter and any significant deviations from the forecast notified to DFID immediately. In addition the Evaluator will be required to provide annual audited statements for the duration of the contract.

Contractual Arrangements

The contract starts in September 2014 and will run till end of December 2019 subject to satisfactory performance as determined through DFID's Annual Review process. Progression is subject to the outcome of this review, strong performance and agreement to any revised work plans or budgets (if revisions are deemed appropriate).

A formal break clause in the contract is included at the end of the inception period. Progression to the implementation phase will be dependent on strong performance by the Evaluator during the inception period and delivery of all inception outputs, including a revised proposal for implementation period. Costs for implementation are expected to remain in line with what has been agreed upon for this contract, with costs such as fee rates fixed for contract duration. DFID reserves the right to terminate the contract after the inception phase if it cannot reach agreement on the activities, staffing, budget and timelines for the implementation phase.

DFID reserves the right to scale back or discontinue this assignment at any point (in line with our Terms and Conditions) if it is not achieving the results anticipated. The Evaluator will be remunerated on a milestone payment basis. DFID has agreed an output based payment plan for this contract, where payment will be explicitly linked to the Evaluator's performance and effective delivery of programme outputs as set out in the ToR and approved workplan. The payment plan for the implementation phase will be finalised during the inception period.

Open Access

The Evaluator will comply with DFID's Enhanced and Open Access Policy. Where appropriate the costs of complying with out open access policy should be clearly identified within your commercial proposal.

Branding
The public has an expectation and right to know what is funded with public money. It is expected that all research outputs will acknowledge DFID support in a way that is clear, explicit and which fully complies with DFID Branding Guidance. This will include ensuring that all publications acknowledge DFID's support. If press releases on work which arises wholly or mainly from the project are planned this should be in collaboration with DFID's Communications Department.

Duty of Care
The Evaluator is responsible for the safety and well-being of their Personnel (as defined in Section 2 of the Contract) and Third Parties affected by their activities under this contract, including appropriate security arrangements. The Evaluator is responsible for the provision of suitable security arrangements for their domestic and business property. DFID will share available information with the Evaluator on security status and developments in-country where appropriate.

The Evaluator is responsible for ensuring appropriate safety and security briefings for all of their Personnel working under this contract and ensuring that their Personnel register and receive briefing as outlined above. Travel advice is also available on the FCO website and the Evaluator must ensure they (and their Personnel) are up to date with the latest position.

The Evaluator has confirmed that:

- The Evaluator fully accepts responsibility for Security and Duty of Care.
- The Evaluator understands the potential risks and have the knowledge and experience to develop an effective risk plan.
- The Evaluator has the capability to manage their Duty of Care responsibilities throughout the life of the contract.

Annex H Index descriptions

1. *Progress out of Poverty Index (PPI) Score:* The Progress out of Poverty Index score uses a poverty measurement tool based on answers to 10 questions about a household's characteristics and asset ownership. The score is then used to compute the likelihood that the household is living below the poverty line. All points in the scorecard are non-negative integers, and total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). The Tanzania index includes questions on the age distribution of household members, school enrolment of children in the household, building materials for the roof and walls, fuel used for cooking, and ownership of certain assets. The latest index was created using Tanzania's 2011/12 Household Budget Survey (HBS). The scorecard is constructed using half of the data from the 2011/12 HBS. That same half of the 2011/12 data is also used to calibrate scores to poverty likelihoods for 18 poverty lines. The other half of the data from the 2011/12 HBS is used to validate the scorecard's accuracy for estimating households' poverty likelihoods, for estimating populations' poverty rates at a point in time, and for segmenting participants.

2. *Assets Index:* An index of household assets was derived from asset and livestock ownership (stove, functioning television, plough, shovel, tractor, bulls, sheep, etc.), using principal component analysis. Scores from the first component, which explained 13% of the total variance, were used to create the total assets index variable. The first component of the consumer durable assets explained 22% of the total variance and was used to create the consumer durable assets index variable; scores from the first component of the production assets, which explained 14% of total variance, were used to create the production assets index; and scores from the livestock assets, which explained 48% of total variance were used to create a livestock assets index.