FANTA III

FOOD AND NUTRITION TECHNICAL ASSISTANCE

USAID FROM THE AMERICAN PEOPLE

Strengthening and Evaluating the Preventing Malnutrition in Children under 2 Years of Age Approach
Burundi Follow-Up Report: Children 24–41 Months

Jef L Leroy
Celeste Sununtnasuk
Jessica Heckert
Deanna Olney

September 2017

FANTA
FHI 360
1825 Connecticut Ave., NW
Washington, DC 20009-5721
Tel: 202-884-8000 Fax: 202-884-8432
fantamail@fhi360.org www.fantaproject.org

fhi 360 THE SCIENCE OF IMPROVING LIVES

This report is made possible by the generous support of the American people through the support of the Office of Health, Infectious Diseases, and Nutrition, Bureau for Global Health, U.S. Agency for International Development (USAID) and the Office of Food for Peace, Bureau for Democracy, Conflict and Humanitarian Assistance, under terms of Cooperative Agreement No. AID-OAA-A-12-00005, through the Food and Nutrition Technical Assistance III Project (FANTA), managed by FHI 360.

September 2017

Recommended Citation

Leroy, Jef L.; Sununtnasuk, Celeste; Heckert, Jessica; and Olney, Deanna. 2017. *Strengthening and Evaluating the Preventing Malnutrition in Children under 2 Approach: Burundi Follow-Up Report for Children 24–41 Months*. Washington, DC: Food and Nutrition Technical Assistance III Project (FANTA).

Contact Information

Food and Nutrition Technical Assistance III Project (FANTA)
FHI 360
1825 Connecticut Avenue, NW
Washington, DC 20009-5721
T 202-884-8000
F 202-884-8432
fantamail@fhi360.org
www.fantaproject.org

Acknowledgments

This publication would not have been possible without the work of the Institut de Statistiques et d'Etudes Economiques du Burundi (ISTEEBU). In particular, the authors would like to thank Nicolas Ndayishimiye, Noé Nduwabike, and Vénérand Nizigiyimana of ISTEEBU. We'd also like to thank Megan Parker, who oversaw the baseline data collection, and Elyse Iruhiriye, who trained, coordinated, and supervised fieldworkers at follow-up and assisted in translations to facilitate data analysis.

Also instrumental in the training process for the baseline survey were the efforts of Leonidas Barihuta and Fortunat Ntafatiro of Programme National Intégré d'Alimentation et de la Nutrition (PRONIANUT), Yvette Mpungenge, and the many families at Rwagasore medical clinic who allowed us to weigh and measure their children during the anthropometric standardization process. For the follow-up survey, we'd like to thank Ampa Dogui Diatta and the community health workers and many families of the Maramvya zone who allowed us to weigh and measure their children during training and the anthropometric standardization process.

We'd like to thank the study supervisors, controllers, household enumerators, anthropometrists, and drivers for their hard work in field conditions that were often difficult.

The surveys were successful in large part because of the tireless efforts of Elisabeth Metellus (baseline) and Rémy Lafalaise (follow-up) in many activities, including questionnaire review, training, and fieldwork supervision. Koen Depaepe was instrumental in writing the Visual Basic for Applications code to create the *sous-colline* specific household listings used by the field teams. Nicole Rosenvaigue provided invaluable language, logistical, and administrative support from the early planning stages of the study to the final steps of data cleaning, analyses, and write-up.

Finally, we express our gratitude to the community members, the health center staff, and the families who graciously gave their time to answer our survey questions, without whom this report would not have been possible.

Contents

List of Tables

List of Figures

Abbreviations and Acronyms

BCC	behavior change communication
BCG	Bacille Calmette-Guérin (tuberculosis vaccine)
BMI	body mass index
cm	centimeter(s)
CRS	Catholic Relief Services
CSB	corn-soy blend
DHS	Demographic and Health Survey
dL	deciliter(s)
DPT	diphtheria, pertussis, and tetanus (vaccine)
DRC	Democratic Republic of the Congo
EICV	Enquête Intégrale sur les Conditions de Vie des Ménages
FANTA	Food and Nutrition Technical Assistance III Project
FFP	Office of Food for Peace
g	gram(s)
HAZ	height-for-age z-score
Hb	hemoglobin
HDDS	Household Dietary Diversity Scale
HFIAS	Household Food Insecurity Access Scale
HepB	hepatitis B virus (vaccine)
HHS	Household Hunger Scale
HiB	*haemophilus* influenza B (vaccine)
HIV	human immunodeficiency virus
IFPRI	International Food Policy Research Institute
IMC	International Medical Corps
INSP	Institut National de Santé Publique
ISTEEBU	Institut de Statistiques et d'Etudes Economiques du Burundi
IU	international unit(s)
kg	kilogram(s)
km	kilometer(s)
km^2	square kilometer(s)
m	meter(s)
mm	millimeter(s)
NGO	nongovernmental organization

ORS	oral rehydration salts
PM2A	Preventing Malnutrition in Children under 2 Approach
SD	standard deviation
SE	standard error
SRQ-20	self-report questionnaire
T18	*Tubaramure* 18
T24	*Tubaramure* 24
THP	*Tubaramure* health promoter
TNFP	*Tubaramure* no food during pregnancy
U.S.	United States
USAID	U.S. Agency for International Development
WHO	World Health Organization

Executive Summary

Tubaramure, a U.S. Agency for International Development (USAID) multi-year development food assistance project, was implemented by a consortium of nongovernmental organizations in eastern Burundi from 2010 to 2014. The project's aim was to improve the health and nutritional status of pregnant and lactating women and children under 2 years of age through three core components: distribution of food rations, participation in behavior change communication (BCC) sessions delivered via care groups, and attendance at preventive health services.

A comprehensive research program conducted by the International Food Policy Research Institute in collaboration with the consortium of nongovernmental organizations implementing *Tubaramure*, and with funding from the Office of Food for Peace at USAID through the Food and Nutrition Technical Assistance III Project (FANTA), was integrated into the *Tubaramure* project. The impact of *Tubaramure* was evaluated via a cluster-randomized controlled trial with three treatment arms. Beneficiaries in the three treatment arms received food rations for varying durations to test the optimal length of provisions: from 4 months gestational age until the child is 24 months old (referred to as the T24 arm), from 4 months gestational age until the child is 18 months old (T18), and from birth until the child is 24 months old (TNFP).

The program's impact was assessed using three cross-sectional surveys (conducted in 2010, 2012, and 2014). The first cross-sectional study provided a baseline reference. The 2012 follow-up survey assessed the impact of *Tubaramure* following approximately 2 years of program implementation on children 0–23 months of age. The corresponding report focused on outcomes such as maternal knowledge regarding health and nutrition-related practices, including infant and young child feeding practices (Leroy et al. 2014). The current report summarizes the findings of the 2014 survey. None of the households in the 2014 follow-up were current beneficiaries, as the program ended 1 month before the follow-up survey started. The objective of the survey was thus to assess the lasting impact of the program after the program ended. We refer to this impact as the *post-program* impact. The 2014 survey focused on (households with) children 24–41 months of age—the group of children exposed to *Tubaramure* consistently from early pregnancy to 24 months of age. The largest program impact on child linear growth (the main measure of nutritional status) was expected to be observed in this age group. The impact on child linear growth in this report thus reflects the full, cumulative effect of the program's impact.

The food component of the *Tubaramure* (rations of corn-soy blend[1] and fortified vegetable oil) was expected to increase household availability of micronutrient-rich food and, in turn, consumption of such foods and to improve diet diversity. The 2012 survey found that participation in food distribution was high among those eligible. *Tubaramure*'s BCC strategy was designed to increase health, hygiene, and nutrition knowledge and to promote the adoption of best practices. The BCC component was implemented through participation in care groups that were directed by leader mothers who were also program beneficiaries. Care group participation, as assessed in the 2012 survey, was low (Leroy et al. 2014). A process evaluation conducted in December 2011 and January 2012 revealed that leader mothers had limited technical expertise and teaching skill and that a large proportion of program beneficiaries were not exposed to the module on complementary feeding since it was rolled out late (Olney et al. 2013). The program's third component focused on health and was designed to improve the provision of preventive and curative health services by health staff (by providing training for health staff and some key

[1] While most FFP development assistance projects switched to CSB+ (which has an upgraded macro- and micronutrient formulation) by 2012, the *Tubaramure* project continued to use the older formulation of CSB to preserve the study's integrity.

supplies) and to increase utilization of these services by pregnant and lactating women and children under 24 months of age.

The 2014 survey found that *Tubaramure* had a large positive post-program effect on household access to food. Small, but significant, positive post-program effects were also found for household and mothers' consumption of a more diverse diet.

The post-program impact on mothers' knowledge was mixed. The program had a positive post-program impact on mothers' knowledge in the area of optimal child feeding practices. *Tubaramure* had a small positive post-program effect on a limited number of household hygiene and sanitation practices. No post-program effect was found on preventive health care utilization. Even after the program had ended, a protective effect on child morbidity was observed from baseline to follow-up, but no post-program effect on curative care-seeking behaviors was observed.

Because of the small number of health centers, no formal impact evaluation of the availability of services and supplies was conducted. Health centers were found to be better equipped and better staffed to care for expecting mothers and young children at follow-up (as compared with baseline). Health education and services available for sick children and children with diarrhea improved dramatically. However, shortages of select immunizations and medications were not uncommon.[2]

An alarming finding was the steep increase in the prevalence of anemia in mothers (by about 12 percentage points) and children (by about 9 percentage points) from 2010 to 2014. The intervention did not have a post-program effect on mean levels of hemoglobin or anemia.

The program had a positive post-program effect on child development: *Tubaramure* was found to have an overall positive post-program effect on the highest language milestone achieved among children 24–29 months; the post-program impact on the highest motor milestone achieved was limited to children 24–29 months in the T24 group.

The second follow-up survey found a steep increase in the already high prevalence of stunting in the control group from 2010 to 2014 (from 68.2 percent to 74.8 percent). *Tubaramure* had a large post-program protective effect on stunting of 5.5 percentage points (all treatment groups combined). The largest effect was found in the T24 group (7.1 percentage points). The effects in the other treatment groups were smaller and did not reach statistical significance. The central research question of the Burundi study was to determine the optimal program duration to improve linear growth. Our findings thus suggest that program exposure from pregnancy to 24 months was necessary for impact on linear growth in the beneficiary population in Burundi. It seems plausible that *Tubaramure*'s effect on improving infant and young child feeding practices (i.e., minimum recommended number of meals, percentage of children receiving the minimum dietary diversity, percentage of children consuming a minimally acceptable diet) and on reducing child morbidity (which we documented in the first follow-up report) (Leroy et al. 2014) contributed to this linear growth impact.

[2] Note that *Tubaramure*, under the USAID Office of Food for Peace guidelines, was not able to purchase immunizations or medications.

1. Introduction

This report presents the findings from the second follow-up survey for the impact evaluation of *Tubaramure*, a Preventing Malnutrition in Children under 2 Approach (PM2A) program, which was implemented in eastern Burundi.[3] This report assesses the impact of the program on children 24–41 months of age — children whose first 1,000 days of life (from conception to 24 months of age, during which they were eligible to participate in the program) fell entirely within the program implementation period.

Tubaramure was a U.S. Agency for International Development (USAID) development food assistance program (formerly called Multi-Year Assistance Program, or MYAP) funded by the Office of Food for Peace (FFP) with Title II resources. The objectives were to improve the health and nutritional status of pregnant and lactating women and children under 2 years of age and to strengthen the quality and delivery of health care services. The program had three core components: distribution of family and individual food rations; behavior change communication (BCC) sessions focused on improving health and nutrition-related behaviors; and promotion of preventive and curative health services for pregnant and lactating women and children under 2 years of age. A consortium of nongovernmental organizations (NGOs) led by Catholic Relief Services (CRS) and composed of the International Medical Corps (IMC), Food for the Hungry, and CARITAS Burundi implemented the program.

A comprehensive research program conducted by the International Food Policy Research Institute (IFPRI) in collaboration with the consortium of NGOs and with funding from the USAID FFP through the Food and Nutrition Technical Assistance Project (FANTA) was integrated into the *Tubaramure* program. The overall objective of the research was to assess the impact and cost-effectiveness of *Tubaramure* on child nutritional status, as well as to evaluate the differential and absolute impact of varying the duration of receiving food rations. The study also sought to evaluate the impact of the program on a number of other household, maternal, and child outcomes, such as household hunger, infant and young child feeding practices, health-seeking practices, maternal levels of hemoglobin (Hb) and anemia, children's morbidity symptoms, children's levels of Hb and anemia, and children's language and motor development. Extensive process evaluation research was also conducted to understand program delivery and utilization (Olney et al. 2013).

In total, three cross-sectional surveys conducted from October to December in 2010, 2012, and 2014 were used to assess the program's impact. The first cross-sectional study provided a baseline reference and is summarized in a prior report (Parker et al. 2012). The second cross-sectional study evaluated the programmatic impact on children 0–23 months[4] (Leroy et al. 2014). The current report uses the results of the third cross-sectional study, which evaluated the impact in children 24–41 months and their mothers.

---

[3] A study conducted in Haiti and funded by USAID through the Food and Nutrition Technical Assistance Project (FANTA) was the first rigorous evaluation conducted under real programmatic conditions that showed that the blanket targeting of a food-assisted maternal and child health and nutrition program to all children 6–24 months of age (preventive approach) was more effective at reducing the community prevalence of stunting, wasting, and being underweight than the traditional approach based on targeting underweight children (weight-for-age z-score < -2) (recuperative approach). Based on the evidence from Haiti, the USAID Office of Food for Peace (FFP) invited proposals to replicate the preventive approach (PM2A) in two other countries: Guatemala and Burundi. The two countries were selected because of their excessively high levels of child stunting. FANTA considered that it would be important to incorporate a strong action-oriented research and development program linked to the implementation of PM2A in the two countries, to allow learning and refinement of the approach and to generate lessons learned for future PM2A programming.

[4] For ease of reference, we use "0–23 months" for children 0–23.9 months of age and "24–41 months" for children 24–41.9 months of age in the remainder of the text.

The remainder of this report is structured as follows. Section 2 presents the study methods. Health center and *colline*[5] characteristics are presented in Section 3. The household, maternal, and child characteristics are presented in Sections 4, 5, and 6, respectively. Section 7 presents information on past *Tubaramure* participation. Sections 8, 9, and 10 follow with results on programmatic impact at the household, mother, and child levels, respectively. Section 11 concludes with a discussion of the results.

[5] *Colline* is the smallest administrative subdivision in Burundi.

2. Methods

2.1 Study Setting

Burundi is one of the 10 poorest countries in the world (United Nations Development Programme 2015). The country is located in the Great Lakes region of East Africa, bordered by the Democratic Republic of the Congo (DRC), Rwanda, and Tanzania. With approximately 10.8 million people living on 25,680 km^2 of land, Burundi is one of the most densely populated African countries (World Bank 2016). Approximately 84 percent of the population lives below the international poverty line of US$1.90 per day (World Bank 2016).

Eighty-eight percent of the population lives in rural areas and more than 90 percent of the labor force is employed in agriculture (Central Intelligence Agency 2016). The majority of rural households depend almost entirely on small-scale subsistence farming (International Fund for Agricultural Development 2012). Rapid population growth (estimated at 3.3 percent in 2014) has continued to fragment household farming areas, resulting in a decline in average household holdings from 0.7 hectares in 1989 to 0.5 hectares in 2009 (International Monetary Fund 2012; World Bank 2016). The most commonly produced food crops are bananas, sweet potatoes, cassava, beans, and maize. Recurrent droughts, declining soil fertility, and insufficient investment in the agricultural sector have contributed to declines in agricultural productivity and increased food insecurity for rural populations (International Monetary Fund 2012).

The *Tubaramure* program was implemented in the eastern provinces of Cankuzo and Ruyigi, located along the border with Tanzania. These provinces are among the poorest of Burundi's 17 provinces and are predominantly rural. According to the 2010 Demographic and Health Survey (DHS),[6] the central-eastern provinces have the greatest prevalence of stunting (62 percent) and being underweight (33 percent) (Institut de Statistiques et d'Etudes Economiques du Burundi [ISTEEBU] and l'Institut National de Santé Publique [INSP] 2011). Micronutrient deficiencies are common: anemia prevalence in these provinces is approximately 44 percent among children 6–59 months and 18 percent among pregnant women; in 2005, more than one-quarter of children 6–59 months and 12 percent of pregnant women were vitamin A deficient (serum retinol <0.7 µmol/l) despite large-scale distributions of vitamin A supplements by the Expanded Program of Vaccination (World Health Organization [WHO] 2009; ISTEEBU and INSP 2011).

2.2 The *Tubaramure* Program

Women were invited to enroll in the *Tubaramure* program when they were pregnant (usually at or after the fourth month of gestation[7]). *Tubaramure* included three core components[8]: receipt of food rations, attendance at preventive health services, and participation in a BCC strategy. The program components are described in more detail below.

Food component.[9] The food component of the *Tubaramure* program was expected to increase household availability of micronutrient-rich food and, in turn, consumption of such foods and improved dietary diversity. To achieve these goals, all *Tubaramure* beneficiaries received a monthly household food ration

[6] Results for the more recent 2015 DHS are not yet publically available.

[7] At the beginning of the program, women could enroll in the program as soon as they found out they were pregnant (if they had proof from a pregnancy test from a health center). However, urine sharing became common. The standard was changed to enroll women into the program if they had proof from a pregnancy test from a health center and if they were showing, which was usually around the fourth month of pregnancy.

[8] FFP limited the focus of the project to the three main areas and did not fund such interventions as agricultural production, livelihoods/income generation, home gardens, and savings/lending groups.

[9] Using private funds, CRS later added a home gardening and animal-raising component. Since these activities were not part of the *Tubaramure* program at the time of the baseline survey, data on these outcomes were not collected.

composed of 12 kg of corn-soy blend (CSB) and 1,200 g of fortified vegetable oil (see **Appendix 1**). This ration was distributed to the beneficiary, but intended for household consumption. In addition to the household ration, an individual ration was distributed for the beneficiary: a pregnant or lactating woman (0–5 months postpartum) received 6 kg of CSB and 600 g of oil, and a child 6–23 months received 3 kg of CSB and 300 g of oil. From the time of enrollment up to the moment the child reached 6 months of age, the beneficiary mother received the monthly individual ration. When the child was 6 months (the age at which it is recommended to introduce complementary foods), the mother's individual ration was discontinued and the child began receiving a monthly individual ration until she or he graduated from the program (at 24 months of age).

Health component.[10] The health component was designed to improve the provision of preventive and curative health services at local health centers and to increase utilization of these services by pregnant and lactating women and by children 0–23 months of age. The improved provision and increased utilization of these services were expected to contribute to improvements in maternal and child health outcomes. The *Tubaramure* program strengthened existing health services through training health staff[11], as well as by providing key supplies for implementing health services. Health supplies included equipment for prenatal care (e.g., vaginal speculum, *Pinard* obstetric stethoscope), labor and delivery (e.g., delivery table), growth monitoring (e.g., salter scale, infant scale), and curative care (e.g., thermometer, sterile equipment). In addition, the *Tubaramure* program through the BCC strategy scri the utilization of preventive and curative health services by pregnant and lactating women (pre- and postnatal services, respectively) and children 0–23 months (growth monitoring and promotion). Beneficiary mothers and children were expected to attend and use recommended preventive health services at the local health center. These included pre- and postnatal check-ups for women and growth monitoring and vaccinations for children.

Behavior change communication component. *Tubaramure*'s BCC strategy encouraged the adoption of best practices in health, hygiene, and nutrition and was designed specifically to address many of the underlying causes of undernutrition in Burundi. The BCC strategy was implemented by CRS and Food for the Hungry staff members and locally hired *Tubaramure* health promoters (THPs). Groups of leader mothers, who were program beneficiaries selected by fellow beneficiary mothers, were first trained by a THP during leader mother care groups. The leader mother care groups were held every 2 weeks. The leader mother in turn taught her fellow beneficiary mothers on the health, hygiene, and nutrition topics that she had most recently learned from the THP; she did this during beneficiary mother care groups, which also met every 2 weeks. The curriculum contained five modules, each with between 6 and 12 lessons.

---

[10] The program followed USAID's FFP guidelines regarding admissible health expenses as per FFP Responses to PM2A Guidelines Questions and Comments January 22, 2009 which stated "Admissible expenses are the same as those listed in the Food for Peace Act Title II Guidance. FFP expects programs to leverage resources from third parties to complement the resources provided for the PM2A that cannot be purchased using Title II resources by linking with existing health services. Please review the new FFPIB 08-03: Eligible Uses of Section 202e and ITSH funding budget matrix found at http://www.usaid.gov/our_work/humanitarian_assistance/ffp/fy09_final_guidelines.html"

[11] IMC provincial technical assistants provided medical training for two nurses and two nurses' assistants working within each district health center and two doctors working within each district hospital. IMC largely focused its efforts on training nurses and paramedics working in the health centers because they were the frontline staff for each community. They received annual training on prenatal care, postnatal care, growth monitoring and promotion, integrated management of childhood illness, and community management of acute malnutrition. IMC also trained the provincial and district health officers to monitor health care activities within the health centers and hospitals to ensure that services were implemented according to the components of the IMC training.

- Module 1 "Care Group Orientation"—six lessons on the program's objectives, teaching techniques, leader mother responsibilities, watching for change and monitoring groups, the value of children, and the ability to change.
- Module 2 "Essential Nutrition, Hygiene, and Care Practices during Pregnancy"—nine lessons on antenatal care services and developing a birthing plan, maternal nutrition, micronutrients (iron and iodine), handwashing with soap (or ash), creation of household handwashing stations, malaria prevention, preparing for delivery and birth, immediate breastfeeding and the use of colostrum, and newborn care practices.
- Module 3 "Essential Nutrition, Hygiene, and Care Practices during Infancy" (first part)—12 lessons on the importance of postpartum care, various aspects of exclusive breastfeeding, childhood illness danger signs, overcoming breastfeeding problems, growth monitoring and promotion, men's involvement in breastfeeding and child care, child spacing, point-of-use water treatment and safe water sources, proper disposal of feces, malaria transmission and prevention, malarial danger signs and treatment, and home care.
- Module 4 "Essential Nutrition, Hygiene, and Care Practices during Infancy" (second part)—seven lessons largely focused on nutrition, including complementary feeding for children 6–8 months of age, 9–11 months of age, and 12–23 months age; preparing CSB with local foods; the importance of vitamin A and good food sources of vitamin A; worms and deworming medication; and preparing, cooking, and storing foods.
- Module 5 "Management of Childhood Infections"—six lessons related to the signs and dangers of child dehydration, dehydration prevention using oral rehydration salts (ORS), proper feeding of sick children, dysentery and persistent diarrhea, pneumonia prevention, and developing a kitchen garden.

In addition to the *collines* participating in the evaluation study (see next section), the *Tubaramure* program was implemented across all *collines* in Cankuzo and Ruyigi. Enrollment in the *Tubaramure* program in the non-study *collines* started in June 2010. The enrollment of new beneficiaries ended in October 2011. Program enrollment in the study *collines* did not start until January 2011 (i.e., after the baseline survey, see below) and continued until October 2012. The provision of program benefits continued in the non-study *collines* until April 2014 and in the study *collines* until September 2014.

2.3 *Tubaramure* Evaluation Design and Study Questions

For the purposes of the IFPRI-led research on *Tubaramure*, 60 study *collines* were randomly assigned to one of four study arms, which differed primarily by timing and duration of food rations, to test the optimal length of provisions provided by the PM2A program:

- *Tubaramure* 24 (T24): the full *Tubaramure* program, including BCC, preventive and curative health services, and food rations (individual and family) during pregnancy and lactation for the mother and for her child up to age 24 months.
- *Tubaramure* 18 (T18): the full *Tubaramure* program, including BCC, preventive and curative health services, and food rations (individual and family) during pregnancy and lactation for the mother and for her child up to age 18 months.
- *Tubaramure* no food during pregnancy (TNFP): the full *Tubaramure* program, including BCC, preventive and curative health services, and food rations (individual and family) starting after the birth of the child, for the mother and for her child up to age 24 months.
- Control: the *Tubaramure* program not provided to these households; however, households could still access preventive and curative services.

The *Tubaramure* health strengthening activities (training for health staff and the provision of key supplies for preventive and curative health services) were implemented in all health centers throughout Cankuzo and Ruyigi, regardless of whether they were located in one of the treatment or control *collines*.[12] These activities may thus have had a positive impact on health-related outcomes in all study arms, including the control group.

The proposed comparison groups and how they relate to the specific research questions are shown in **Table 1**.

Table 1. Study arm comparisons and research questions

Study arm	Compared to	Research question
T24	Control	What is the impact of *Tubaramure* (compared to a control group) on child nutritional status?
TNFP	Control	What is the impact of *Tubaramure without food rations during pregnancy* (individual + family) on child nutritional status?
T18	Control	What is the impact of receiving *Tubaramure food rations only up to the age of 18* months on child nutritional status?
TNFP	*T24*	What is the differential impact of *Tubaramure without food rations during pregnancy* (individual + family) vs. *Tubaramure*?
T18	*T24*	What is the differential impact of receiving *Tubaramure* up to the age of 18 months *vs.* receiving it up to the age of 24 months?

All research procedures involving human subjects were approved by IFPRI's institutional review board, and the study was approved by the Ministry of Health in Burundi. Verbal informed consent was obtained from the primary household respondent before the start of each interview at baseline and follow-up.

2.4 Study Methods

The main outcome of the study was child nutritional status. The largest program impact on child linear growth (the main measure of nutritional status) was expected to be observed in children exposed to *Tubaramure* consistently from early in-utero to 23 months of age. The first group of children to meet this condition were 24–41 months of age in October to December 2014 (i.e., 4 years after the baseline survey, carried out from October to December 2010), when the survey for this report was conducted. None of the households in this survey were current beneficiaries, as the program had ended 1 month before the start of the survey. The impact on child linear growth in this report thus reflects the full, cumulative effect of the program's impact. The impact on other outcomes (such as knowledge, health practices, and dietary diversity) reflect the lasting impact of the program after the program ended. We refer to this impact in this report as the post-program impact.

For the short-term or immediate impact on these outcomes, we refer readers to the report of the 2012 follow-up survey (Leroy et al. 2014). To reduce recall bias, these outcomes were assessed among mothers of children 0–23 months of age in 2012, when the program was still being implemented.

2.4.1 Sample Size

Sample size calculations for children 24–41 months of age were based on the estimated program impact on height-for-age z-score (HAZ) and stunting using the differences expected to be found at follow-up between the three intervention groups and the control (first three study questions in Table 1), between

[12] Not every *colline* has a health center. (Of the 60 *collines* in the study sample, only 13 had a health center.) As a consequence, the health strengthening activities could not be randomized.

T24 and T18 (fourth research question), and between T24 and TNFP (fifth research question). One-sided test assumptions were used when clear *a priori* assumptions about the direction of the effect were justified. The required sample size in each study arm is shown below (**Table 2**). The sample size was calculated with the following parameters: 0.05 probability of a type I error, power of 0.90, and intra-cluster correlations of 0.006 for stunting and 0.009 for HAZ.

Table 2. Required sample size to measure impact on height-for-age z-score

	Study arm			
	T24	T18	TNFP	Control[13]
Age (months)	24–41	24–41	24–41	24–41
Minimum sample size required (by age)	1,000	1,000	1,000	584

2.4.2 Sampling

To systematically select the research *collines*, the 210 *collines* meeting certain population and primary health care provision criteria were ranked according to population size and divided into five strata (based on population size) in Cankuzo (13 or 14 *collines* per stratum) and 10 strata in Ruyigi (14 or 15 *collines* per stratum). The number of strata per province was based on the relative population size. Within each stratum, four *collines* were selected using random numbers with a fixed random number seed in Stata version 11 (StataCorp 2009). The four *collines* in each stratum were then each assigned randomly to one of the four study arms in a public event organized for both study provinces in the administrative center of Ruyigi on January 25, 2010. The list of selected *collines* and how they were assigned to the study arms is shown in **Appendix 2**. The random allocation of study *collines* is shown in **Figure 1**.

A household census was conducted in the 60 research *collines* to generate a complete list of households (with children under 60 months of age at baseline and 12–60 months of age at follow-up) before the start of both the baseline and follow-up surveys (October of 2010 and 2014, respectively). An enumerator visited each household in the 60 research *collines* to inquire if there was a child meeting the age requirement. The following information was recorded on a pre-numbered census sheet if the household met the requirement: the *sous-colline*[14] where the household lived, the first and last names of the head of household, and the household size. The head of the household was given a pre-printed slip with a unique number corresponding to the number on the census sheet. After all households were identified and recorded, the data were entered and transferred to IFPRI.

Using the household census data, the target sample sizes for children 24–41 months of age (baseline and 2014 follow-up) were calculated for each *sous-colline*. At baseline, the target sample size for each *sous-colline* was calculated by first dividing the total number of households with a child under 5 years in each *sous-colline* by the total number of households with a child under 5 years across the 15 *collines* of the respective study arm. The target sample size was then calculated for the 24–41 month age group in each *sous-colline* by multiplying the *sous-colline*-specific proportion by the arm- and age group-specific sample size needed. The same procedure was used at follow-up to calculate the number of children 24–41 months to be sampled.

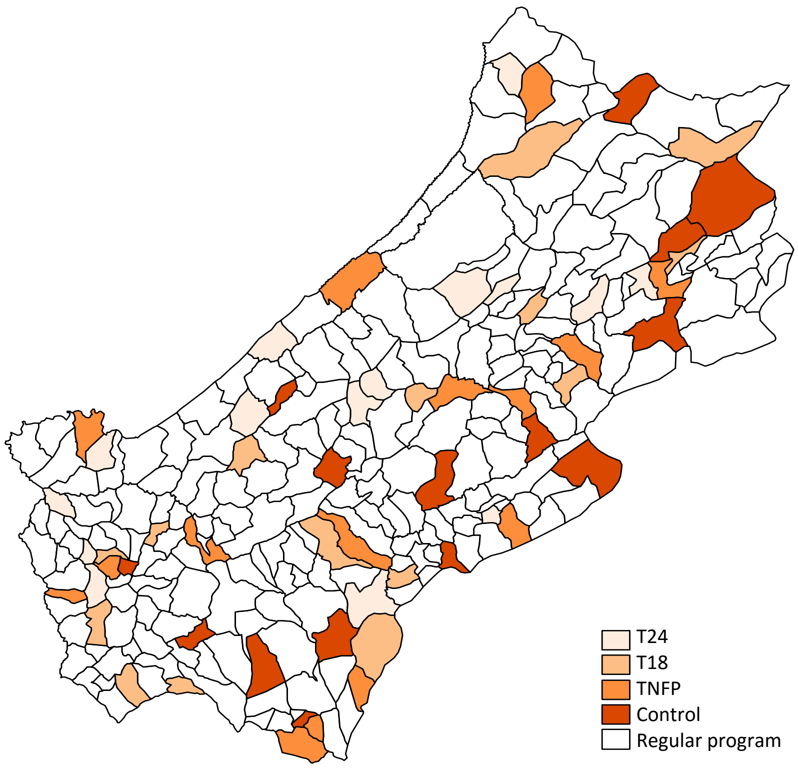
A list of the potential households to be surveyed was generated for the field team in each of the *sous-collines.* Each *sous-colline*-specific list showed the location identifiers and the total number of households

[13] The smaller sample size requirement in the control group was a consequence of the fact that the expected difference between the treatment arms and the control was larger than that between the different treatment arms.

[14] Each *colline* is divided into a number of *sous-collines*.

to be visited. Each list further showed the (randomly ordered) households with a child potentially meeting the age criteria. The supervisor of each field team was instructed to visit households in the order that they were listed until the required sample size was reached. Households were revisited when the respondent was not available.

Figure 1. Allocation of *collines* to study arms in Ruyigi and Cankuzo



2.4.3 Selection of Index Child

At baseline, if there was only one child under 42 months of age in the household, that child was designated as the "index child." If there was more than one child under 42 months of age, the children's first names were alphabetized, and the first child on the list was selected as the index child.[15] At the 2014 follow-up, households were included in the sample based on having a child 24–41 months of age. If there was more than one child 24–41 months of age, the children's first names were alphabetized, and the first child on the list was selected as the index child.[16]

2.4.4 Selection of *Colline* Leaders and Health Centers

Data were collected from a group of leaders in each *colline* and from all public health centers located in a research *colline*. Several community leaders in each *colline* were invited to meet to complete the *colline*-level questionnaire. A survey team member helped the group collaboratively complete the *colline*-level questionnaire. The *colline*-level interviews were conducted at both baseline and the 2014 follow-up. After

[15] At baseline, it was determined that mothers with two children 0–41 months of age would be interviewed only about one child, for two reasons. First, we wanted to reduce the respondent burden to the extent possible. Second, lengthy surveys can lead to poorer data quality (United Nations Department of Economic and Social Affairs Statistics Division 2005).

[16] The prevalence of multiple children 24–41 months of age was low (2 percent). No weights were used in the analyses.

the completion of *colline*-level interviews at the 2014 follow-up, questionnaires for six *collines* were stolen in the field. Thus, data for only 54 *collines* are presented at follow-up.

All public health centers located within each of the 60 research *collines* were invited to be interviewed. At baseline, the study team identified 13 public health centers in the research *collines*. A member of the anthropometry team interviewed a member (or multiple members) of the staff capable of completing the survey at each center. It was later determined that one of the putative health centers was actually a hospital, and it was therefore not revisited at follow-up. Additionally, personnel from one center did not agree to an interview at follow-up. Thus, the description of health centers is based on the 11 health centers for which we have both baseline and follow-up data.

The sample sizes for the evaluation were calculated to detect meaningful changes in child HAZ measured at the individual level. The study was not powered to detect meaningful changes at the health center or *colline* level. Therefore, these data are mainly intended to be descriptive.

2.4.5 Data Collection

Data were collected at the *colline*, health center, and household levels, using pretested questionnaires. The list of modules included within each of the questionnaires, as well as brief module descriptions, are presented in **Tables 3**, **4**, **5,** and **6**.

Colline Questionnaire

The *colline* questionnaire collected information on the availability of and access to schools and health services, food crops, and fruit trees; the presence of associations or cooperatives; forms of transportation; infrastructure; recent immigration/emigration patterns; weather conditions; development projects; and positive and negative events that affected the *colline* residents.

Table 3. List of modules included in *colline* questionnaire

Module	Topic	Description	Respondent
1	Schools	Identified schools attended by children living in the *colline*, including location, type, costs, and perceived quality	Group of community members
2	Health services	Identified health services used by families living in the *colline*, including location, travel time, vaccination campaigns, and epidemics	Group of community members
3	Food crops	Identified the main crops in the *colline* and timing of harvest	Group of community members
4	Fruit trees	Identified the main fruit trees in the *colline* and timing of harvest	Group of community members
5	Community organizations	Identified existing organizations, objectives, and membership	Group of community members
6	Transportation	Investigated the distance from and the availability and cost of public transportation to various public services	Group of community members
7	Infrastructure	Investigated the availability of electricity, water, and telephone services	Group of community members
8	History	Investigated migration, climatic conditions, and living conditions over the past 5 years	Group of community members
9	Development programs	Identified the presence and outcomes of local development programs over the past 5 years	Group of community members
10	Events	Investigated local events (positive, negative) that affected the *colline* over the past 5 years	Group of community members

Health Center Questionnaire

The health center questionnaire gathered information on the health center's schedule, personnel, services provided, vaccinations provided, services for children and mothers provided, equipment, and medical supplies. Differences between the baseline and follow-up questionnaires are identified in **Table 4**.

Table 4. List of modules included in health center questionnaire

Module	Topic	Description	Respondent
1	Schedule	Identified the hours of operation for emergency services, prevention services, prevention services for pregnant women, and prevention services for postpartum women	Health center personnel
2	Personnel	Identified the number of personnel per establishment, personnel qualifications, and the services to which the personnel attend	Health center personnel
3	Services	Identified the health services provided by each establishment for women and children *Differences between baseline and follow-up survey modules:* • The follow-up survey also included monitoring growth of children under 5 years via mobile services (i.e., facility staff travel to another place to provide services)	Health center personnel
4	Vaccinations	Evaluated the availability of vaccines and vitamin A *Differences between baseline and follow-up survey modules:* • The follow-up survey also included the PCV 13 (pneumococcal conjugate), VAROTA (rotavirus), and RR (measles and rubella) vaccines	Health center personnel
5	Services for children	Investigated the provision of growth monitoring services, examinations for sick children, and treatment for severely malnourished children	Health center personnel
6	Services for women	Investigated the provision of prenatal care, delivery assistance, and postnatal care *Differences between baseline and follow-up survey modules:* • The follow-up survey also included praziquantel (for *taenia* deworming) and tests for malaria under prenatal services provided to women	Health center personnel
7	Equipment	Evaluated the presence of equipment and material available for preventive and curative care for children and pregnant women *Differences between baseline and follow-up survey modules:* • At follow-up, types of registries (asked only as registries at baseline) were subdivided into curative consultations, growth monitoring, vaccinations, and prenatal care • Immunization cards (at baseline) were further specified as maternal health cards, child health cards, and child health passports at follow-up • Disposable needles and syringes were a single category at baseline and listed separately at follow-up • Additional items at follow-up included wheelchair, trash can, sterile dressings, sink, soap, towels/napkins, metal tongue depressor, flashlight, a source of suction, and packet of urine sample dipsticks	Health center personnel
8	Medications	Evaluated the presence of medications and ruptures in their supply *Differences between baseline and follow-up survey modules:* • Co-trimoxazole and erythromycin, as well as benzylpenicillin and gentamycin, were listed together at baseline and separately at follow-up. • Additional items at follow-up included tetracycline ophthalmic ointment (for measles treatment), nystatin, praziquantel (for *schistosoma* and *taenia* treatment), diclofenac sodium, aspirin, ibuprofen, niclosamide, oral contraceptives, multivitamin pills, drinkable multivitamins, vitamin B complex, and antiretroviral medications (including emtricitabine, lamivudine, zidovudine, tenofovir, and nevirapine)	Health center personnel

Household Questionnaire

The household questionnaire gathered information on household, maternal,[17] and child characteristics. The household questionnaire was based on the questions used in the Rwanda DHS

[17] Data were collected on the mother or caregiver of the index child. Where the index child did not live with his/her biological mother, data were collected from the caregiver of that index child.

(www.measuredhs.com); the Rwanda Enquête Intégrale sur les Conditions de Vie des Ménages (EICV 1998); and other instruments from a variety of sources, including the FANTA Household Hunger Scale (HHS), the FANTA Household Food Insecurity Access Scale (HFIAS), and the FANTA Household Dietary Diversity Scale (HDDS) (Coates et al. 2007; Deitchler et al. 2010; Swindale and Bilinsky 2006). All modules were adapted to the specific needs of this study. **Table 5** presents the modules included in the questionnaire, the questionnaire or instrument the module was based on, and a short description of each module. We also indicate where the baseline and follow-up questionnaires differed.

Table 5. List of modules included in household questionnaire

Module	Topic	Source	Description	Respondent
1	Household roster and education	IFPRI	Collected information on the composition of the household, including designation of the head of household, a list of all household members, their ages* and sex, and their relationship to the head of household, (biological) parents of the children under 5 years of age, and the highest educational level attained and activity/employment in the past month of all household members at least 3 years of age * Each child's birth date was obtained from the birth certificate, from the child's vaccination card, or from recall if neither document was available	Head of household, spouse, or household member over 18 years of age
2	Eligible child	IFPRI	Identified all children 24–41 months of age and randomly selected the index child	Head of household, spouse, or household member over 18 years of age
3	Dwelling	EICV	Identified construction materials used for floor, walls, and roof; availability of water and electricity; fuel/energy used for cooking, lighting, etc.	Mother or household member over 18 years of age
4	Assets	EICV, IFPRI	Identified ownership of durable household goods (in working condition), including tools for agricultural production	Mother or household member over 18 years of age
5	Household dietary diversity	FANTA HDDS	Evaluated the diversity of the household diet in the past 24 hours	Individual in charge of food preparation, or household member over 18 years of age
6	Participation in social programs	IFPRI	Identified all social programs that household members participate in and the benefits received from those programs *Differences between baseline and follow-up survey modules:* • A section on past participation in *Tubaramure* was added at follow-up	Head of household, spouse, or household member over 18 years of age
7	Shocks	IFPRI	Identified all shocks (economical, agricultural, and familial) faced by the household in the past 12 months	Head of household, spouse, or household member over 18 years of age
8	Food security	FANTA HHS, HFIAS	Investigated the prevalence of household hunger using the FANTA HHS *Differences between baseline and follow-up survey modules:* • At follow-up, HFIAS was used to measure the access component of household food insecurity; the HHS was calculated from the HFIAS collected at follow-up	Individual in charge of food preparation or household member over 18 years of age

Module	Topic	Source	Description	Respondent
9	Maternal knowledge	DHS, IFPRI	Evaluated knowledge on child health, health care seeking, child feeding practices, and danger signs during pregnancy *Differences between baseline and follow-up survey modules:* • Knowledge section was revised based on the content of *Tubaramure*'s BCC sessions • Foods essential for child growth were not asked at follow-up	Mother of index child
10	Women's status	DHS, IFPRI	Evaluated women's empowerment and decision-making power	Mother of index child
11	Women's occupation and activity	IFPRI	Evaluated maternal literacy, occupation, and activities	Mother of index child
13[18]	Maternal health	IFPRI, SRQ-20	Evaluated maternal health and stress *Differences between baseline and follow-up survey modules:* • A section on the mother's dietary diversity was added at follow-up	Mother of index child
14	Feeding practices	IFPRI	Investigated the child's dietary diversity during the past 24 hours	Mother of index child
15	Child health	WHO, integrated management of childhood illness, IFPRI	Evaluated preventive health care utilization, vaccination status, morbidity, and curative health care seeking *Differences between baseline and follow-up survey modules:* • The follow-up survey also included the status of measles vaccination	Mother of index child
16	Child development	Multiple Indicator Cluster Surveys — Round 4	Evaluated the presence of books and toys for children, child care, and child development	Mother of index child
17	Hygiene spot-check	IFPRI	Evaluated the cleanliness of the child, mother, and interior and exterior of the house	Enumerator (direct observation)

The anthropometry questionnaire was used to record the height, weight, and Hb concentrations of index children and their mothers (**Table 6**).

[18] Section 12 was excluded, since it applied to children 0–23 months of age.

Table 6. List of modules included in anthropometry questionnaire

Module	Topic	Description	Respondent
18	Child anthropometry	Child weight and length/height were measured; length or height was measured twice and a third time if the difference between the first two measurements exceeded 6 mm;[19] weight was measured once	Index child
	Maternal anthropometry	Mother's height and weight were measured and pregnancy status recorded; maternal height was measured twice and a third time if the difference between the first two measurements exceeded 1 cm;[20] weight was measured once	Mother and index child
	Mother and child Hb concentrations	Hb concentrations of the index child and his/her mother were measured	Mother and index child

2.4.6 Field Work

Census

The firm ISTEEBU hired and trained data collectors with previous experience conducting censuses. The training included an orientation to Cankuzo and Ruyigi provinces, organization into field teams, and instructions to visit each house within the 60 *collines*. IFPRI provided preprinted census sheets for each *sous-colline* to record the name of the head of the household and the number of household members for each household with a child meeting the age criterion.

Colline, Health Center, and Household Surveys

- **Survey firm.** ISTEEBU was selected to conduct the baseline and follow-up surveys.
- **Survey teams.** Each survey team (22 at baseline, 8 at the 2014 follow-up) was composed of four enumerators, two anthropometrists, and one team controller. The enumerators conducted the household survey, the anthropometrists collected and recorded the anthropometric and Hb data and conducted the health center survey, and the controllers checked the quality of the data and conducted the *colline* surveys. The supervisors (four at baseline, two at follow-up) oversaw the controllers and reported to the coordinating team based in Bujumbura. Survey teams were assigned to specific *collines* and were monitored closely and continuously by the survey firm and by IFPRI staff and consultants throughout the fieldwork. Continuous monitoring ensured that challenges encountered during fieldwork were addressed in an efficient and timely manner and that a high level of data quality was maintained throughout the survey.
- **Training of supervisors.** The supervisors helped pretest the French version of the questionnaire and verified and revised the Kirundi version as needed. Therefore, they were well oriented to the project prior to the training of controllers, enumerators, and anthropometrists. The supervisors attended both the 1-week training for controllers and the 4-week training for enumerators. They were trained to manage a large team of enumerators, organize the questionnaires, and monitor and report progress and problems to ISTEEBU.
- **Training of controllers.** The controllers completed 1 week of specialized training using lectures, role-plays, discussions of potential responses, and discussions of how to code responses. Following this

[19] If the difference between the first two measurements was less than 6 mm, the child's length/height was calculated by taking the average of the two measurements. If a third measurement was required, the child's length/height was calculated by taking the average of the two measurements with the least difference between them.

[20] If the difference between the first two measurements was less than 1 cm, the mother's height was calculated by taking the average of the two measurements. If a third measurement was required, the mother's height was calculated by taking the average of the two measurements with the least difference between them.

training course, all controllers participated in the enumerator training. Controllers assisted the enumerator training process by leading small group role-plays.

- **Training of enumerators.** A variety of methods (including lectures, role-plays, discussions of potential responses, and discussions of how to code responses) were used to train the enumerators in the use of the survey questionnaire over 4 weeks. The controllers and enumerators were continuously evaluated during the training. A weekly short written test was used to evaluate their understanding of the questionnaire. The final selection of controllers and enumerators was based on a short field-based evaluation at the end of the training.
- **Standardization of the anthropometry team.** The fieldworkers who conducted the anthropometric measurements were carefully trained for 2 weeks in the use of the equipment and the recording of anthropometric data for children and mothers. Training of the anthropometry team made use of lectures, videos, and equipment demonstrations and was followed by practical exercises. The fieldworkers were then standardized (Cogill 2003) in the measurement of height. First, the trainer and all fieldworkers measured the height of a group of children and their mothers twice. A spreadsheet was created to compute the precision and accuracy of all trainees. A second round of standardization was organized for those needing more practice. Based on the results of the standardization, the most accomplished anthropometrists were selected. Height data were collected using Shorr boards (Weigh and Measure, USA). Weight data were collected using a Seca 874 digital scale (Seca, USA).
- **Training on the collection of capillary blood.** The members of the anthropometry team were trained to collect capillary blood from mothers and children using a sterile spring-loaded lancet (which standardizes the depth of the puncture and minimizes the need to puncture more than once). They were also trained on the proper procedure for taking the sample and measuring the Hb concentration using portable HemoCue 201+ analyzers (HemoCue, USA). Training covered proper techniques for sterile sample collection, including how to use a sterile lancet, and making sure a new set of sterile gloves was used for each sample. The enumerator used an alcohol swab to clean the area to be pricked both before and after the procedure and applied a small bandage when finished. Each enumerator was supplied with a small plastic container to collect the waste from the procedure.
- **Administration of the *colline* questionnaire.** Each *colline* questionnaire was conducted using a group interview methodology. The questionnaire was administered by at least two controllers who ensured that a consensus was reached on all responses.
- **Administration of the health center questionnaire.** The anthropometrists administered the health center questionnaire to at least one health professional per health center.
- **Administration of the household and anthropometry questionnaires.** An information session was organized in each *colline* before the survey work began. These sessions explained the purpose of the survey to the community members; provided reasons for measuring children's and mothers' height, weight, and Hb concentrations; and requested the cooperation of community members. Enumerators reviewed each questionnaire before leaving the household where it was administered. Controllers reviewed each questionnaire for accuracy, logical patterns, and legible writing at the end of each day. Enumerators and anthropometrists returned to households when the controller observed missing data or other problems.
- **Handling of questionnaires.** Upon completion of each household interview, the enumerators submitted their questionnaires to their respective controller. The controller recorded the number of questionnaires completed per *sous-colline* to monitor whether the quotum was reached. Once all the questionnaires in a single *colline* were completed, the controller gave the completed package to her or his supervisor. The supervisors then checked the questionnaires they had received from controllers for

quality, and they recorded and reported the total number of questionnaires completed by *colline* to the coordinators. Questionnaires were grouped according to *colline*, commune, and province, and then transported to the ISTEEBU office in Bujumbura for data entry.

2.4.7 Data Entry and Cleaning

Data Entry

ISTEEBU developed a data entry program using CSPro and provided the data entry clerks with an interface resembling the paper questionnaire. Two data entry clerks entered each of the health center, *colline*, household, and anthropometry questionnaires into the CSPro program. At the end of the data entry period, lists of inconsistencies between the first and second entries of the same questionnaire were generated and mistakes corrected.

Data Cleaning

The data were transferred to Stata (StataCorp Stata Statistical Software: Release 14), and IFPRI performed standard data cleaning checks. All observations with problems were listed and verified using the paper questionnaires.

Variable Creation

From the data collected, new variables were created to summarize health center, household, and individual characteristics more concisely. Many of these variables were based on norms and standards provided by international organizations and the government of Burundi. The variables created in this study are summarized below.

Health center variables. We compared the information obtained from the health center questionnaires to the national guidelines for health centers (Ministère de la Santé Publique et de la Lutte contre le SIDA 2012a and 2012b). These guidelines describe the qualifications of essential personnel; the recommended characteristics of the facilities; and the essential equipment, medications, and supplies that each health center should stock.

- **Personnel.** These are indicators for whether the health center staffs the personnel recommended by national guidelines. An A2-level nurse has completed 4 years of nurse training after 10th grade. An A3-level nurse has completed 2 years of training after 10th grade. A nutritionist is an A3-level nurse with specialized nutrition training. An A2-level laboratory technician has completed 4 years of specialized training after 10th grade. (Training requirements for other personnel are not described.)
- **Total number of services for sick children.** This value sums the number of essential services reported to be available to sick children visiting the facility: measure body temperature, measure weight, chart weight, offer antibiotics, offer acetaminophen for fever, provide health education, evaluate immunization status, and evaluate vitamin A status (range of the created variable 0 to 8, with 0 indicating that none of the eight services were available, and 8 that all of them were available).
- **Total number of services for children with diarrhea.** This value sums the availability of four essential services for children with diarrhea: ORS packets, oral rehydration therapy at the health center, zinc supplementation, and onsite intravenous solution in the case of severe diarrhea (range 0 to 4).
- **Total number of prenatal care components.** This value sums the number of essential service components reported to be offered to pregnant women during prenatal visits at the facility: measure weight, measure abdomen, measure blood pressure, offer tetanus vaccine, distribute iron folate or iron and folate, offer education sessions, administer albendazole or mebendazole for deworming, and administer niclosamide for *taenia* deworming (range 0 to 8).

- **Rupture in the supply chain.** A health center experienced a rupture in the availability of essential vaccines or medications if personnel reported that they experienced a shortage of any duration during the past 6 months.

Household characteristics. The following household variables were created.

- **Dependency ratio.** The ratio of economically dependent household members (under 15 years of age or over 60 years) to economically active ones (15–60 years of age).
- **Cleanliness of mothers, children, and interior and exterior of dwellings.** These variables were constructed from spot-check observations conducted at the time of the interviews. Fieldworkers[21] noted the cleanliness of hands, face, hair, and clothes of mothers and children, and noted the presence of garbage, feces, dust, or dirty clothes around dwellings. The variables describe the percentage of people or premises scoring "clean" on all counts.[22]
- **Assets.** Household asset ownership was summarized in six count variables: the total number of household goods, the total number of agricultural tools and equipment, the total number of small animals, the total number of medium-sized animals, the total number of large animals, and the total number of motorbikes or bikes.
- **Household hunger scale.** This scale was constructed according to FANTA guidelines (Ballard et al. 2011; Deitchler et al. 2010), with scores assigned to a set of three questions about meals and hunger ("no food to eat of any kind in your household;" "go to sleep at night hungry;" "go a whole day and night without eating"), based on the frequency of occurrence (never = 0; rarely or sometimes = 1; often = 2) over the past 4 weeks. A total score (range 0 to 6) was calculated and the following classifications made: 0–1, "little or no hunger;" 2–3, "moderate hunger;" and 4–6, "severe hunger."
- **Household dietary diversity score.** According to FANTA guidelines (Swindale and Bilinsky 2006), the food preparer in each household was asked if any member of the household had consumed food from 12 predefined food groups[23] in the past 24 hours. A simple score ranging from 0 to 12 was provided.
- **Household food insecurity access scale.** This scale was constructed according to FANTA guidelines (Coates et al. 2007), with scores assigned to a set of nine questions ("worry household would not have enough food;" "unable to eat preferred foods;" "eat limited variety;" "eat unwanted foods;" "eat smaller meals;" "eat fewer meals;" "no food to eat in your household;" "go to sleep at night hungry;" "go a whole day and night without eating"), based on the frequency of occurrence (never = 0; rarely = 1; sometimes = 2; often = 3) over the past 4 weeks. The sum of these responses is the household score (range 0 to 27). An algorithmic tabulation plan classified households as food secure, mildly food insecure, moderately food insecure, or severely food insecure.

---

[21] Fieldworkers were extensively trained on this instrument, but no formal standardization was conducted.

[22] For mothers, fieldworkers were asked to assess the cleanliness of hands, hair, clothes, and face. Possible answers were "clean," "dirty," or "dusty." The same variables and answers were used for the children. Mothers and children were classified as "clean" if the fieldworker recorded "clean" for all items. The outside of the house was evaluated with respect to the need for cleaning and for the presence of human feces, animal feces, and garbage. The inside of the house was evaluated with respect to the need to be swept, the presence of animal feces, the water stored at home being covered, and the presence of dirty clothes. The outside of the house was classified as "clean" if the fieldworkers recorded "no" for all items. The same approach was followed for the inside of the house.

[23] The 12 HDDS food groups are: cereals and grains; roots and tubers; legumes, nuts, and pulses; milk and dairy products; eggs; meat and poultry; fish and seafood; fruits; vegetables; oils and fats; sugar, honey, sweets, and snacks; and miscellaneous.

Maternal characteristics. The following variables were created to describe maternal characteristics.

- **Maternal literacy.** Literacy was evaluated by asking mothers to read one of two sentences in Kirundi. The women were classified as literate if they could read the entire sentence, partially literate if they could read a little, and illiterate if they could not read the sentence at all.
- **Maternal knowledge.** Mothers were asked a series of questions to assess their knowledge of danger signs for childhood illness; how to care for a sick child or a child recovering from an illness; knowledge of appropriate breastfeeding, complementary feeding, and causes of malnutrition; and optimal hygiene practices for the prevention of diarrhea and worms. Variables were created to describe the percentage of mothers responding correctly to each knowledge question within these four categories.
- **Maternal dietary diversity.** Maternal diet was calculated using a dietary diversity score (Kennedy et al. 2011). Based on the mother's 24-hour dietary recall, all foods and liquids consumed were classified into one of nine food groups (starchy staples, dark green leafy vegetables, vitamin A-rich fruits and vegetables, other fruits and vegetables, organ meat, flesh foods, eggs, legumes/nuts/seeds, and milk and milk products). Data to create these variables were collected only at follow-up.

Preventive health care practices. Preventive health care practices reported by mothers were evaluated in relation to national recommendations, detailed below.

- **Preventive child care practices.** The standards for growth monitoring, vitamin, and vaccination coverage were based on national (or, if not available, international) standards. Information about child weight, supplements, and vaccinations were provided by the child's vaccination card. Information about these preventive care visits was recorded directly from the vaccination card if the primary caregiver presented the card at the time of the interview. If the primary caregiver could not present the vaccination card, she was asked to recall the preventive care that the child had received during the care-specific recall period. The following variables were created.
  - **Growth monitoring visits.** According to *Tubaramure* recommendations, children 24–42 months should be taken to a clinic every 3 months for growth monitoring.[24,25]
  - **Vitamin and mineral supplements.** According to national recommendations,[26] children should begin receiving vitamin A supplements at 9 months and continue receiving vitamin A every 6 months until 5 years of age.
  - **Vaccination.** The national vaccination schedule (**Table 7**) was used to calculate the percentage of children fully immunized. Only information recorded from vaccination cards was used to construct these variables, and vaccination information collected from maternal reports was considered missing.

---

[24] If the *Tubaramure* recommendation was followed, approximately one-third of respondents would be expected to report having taken their child to the clinic for growth monitoring in the past month. A percentage lower than 33 percent indicates that growth monitoring attendance is lower than recommended.

[25] The *Tubaramure* process evaluation conducted in 2011 found that activities were limited to growth monitoring (not growth monitoring and promotion). In most cases, only the weight was recorded, and even when moderately malnourished children were identified, they were not counseled or referred for treatment (Olney et al. 2013).

[26] WHO's current recommendation is to start supplementation at 6 months. The Burundi recommendation is thus different from WHO's.

Table 7. National vaccination schedule for Burundi

Age	Vaccination
Birth (0 weeks)	BCG,* Polio-0
6 weeks	DPT1-HiB1-HepB1,** Polio-1
10 weeks	DPT2-HiB2-HepB2, Polio-2
14 weeks	DPT3-HiB3-HepB3, Polio-3
9 months	Measles

* Bacille Calmette-Guérin (tuberculosis vaccine)
** Vaccines for diphtheria, pertussis, and tetanus (DPT), *haemophilus* influenza B (HiB), and hepatitis B

Anthropometric and hemoglobin measures. Anthropometric data were used to construct the following indicators.

- **Maternal body mass index.** Body mass index (BMI) was calculated for non-pregnant women as weight (kg)/height (m).2 Three BMI categories were created: underweight (BMI < 18.5), normal weight (BMI ≥18.5 and < 25), and overweight/obese (BMI ≥ 25) (WHO Expert Committee on Physical Status 1995).

- **Child height-for-age z-score, weight-for-height z-score (WHZ), and weight-for-age z-score (WAZ)**. These z-scores were calculated using the 2006 WHO growth standard (WHO Multicentre Growth Reference Study Group 2006). Stunting was defined as HAZ <-2 standard deviation (SD), wasting as WHZ <-2 SD, and underweight as WAZ <-2 SD.

- **Maternal and child hemoglobin and anemia.** Hb concentrations vary with altitude. The Hb concentration values were thus adjusted according to international guidelines (Nestel 2002; Stevens et al. 2013; WHO 2011a) using the following formula:

$$measured\ Hb\ + \frac{(0.32*(altitude*0.0033))-(0.22*(altitude*0.0033)^2)}{10}\ if\ altitude > 1{,}000\ \text{m}$$

  Anemia was defined for non-pregnant women as having an Hb concentration < 12 g/dL and for pregnant women and children as having an Hb concentration < 11 g/dL. Severe anemia was defined for non-pregnant women as having an Hb concentration < 8 g/dL and for pregnant women and children as having an Hb concentration < 7 g/dL (WHO 2011a; WHO 2011b).

Child development indicators. Children's motor and language development were assessed by parental report using a predefined list of motor and language milestones ranked in order to reflect a generally accepted sequence of achievement. The motor milestone scale consists of 30 motor milestones ranging from the first milestone of a child being able to hold his or her head straight, to the 30th milestone of skipping using alternate legs. The full list of motor milestones is presented in **Appendix 3**. The language milestone scale consists of 21 milestones ranging from the first milestone of making sounds while playing alone, to the 21st milestone of talking about things that took place in the past. The scales are adapted from ones previously used in Tanzania (Olney et al. 2009; Stoltzfus et al. 2001). Items are ordered to reflect the sequence of motor and language development and are derived from the Griffiths and McCarthy scales (Griffiths 1970; McCarthy 1972). The full list of language milestones is presented in **Appendix 4**.

Parents were asked if their child had achieved each of the motor and language milestones. Once three milestones were recorded as not being achieved, the interviewer stopped asking about the remaining milestones. In addition to the parental report, children were asked to demonstrate key motor milestones, such as crawling, walking with and without assistance, and standing alone (WHO Multicentre Growth Reference Study Group 2006).

The impact of the *Tubaramure* program on child development (motor and language development) was examined by looking at differences between groups from baseline to follow-up on the highest language and motor milestones attained. The highest attained milestone was defined as the highest milestone reported by the parent. Impacts were assessed for the full sample of children 24–41 months of age and then within three age ranges (24–29 months, 30–35 months, and 36–41 months).[27]

2.4.8 Data Analysis and Impact Estimation

Descriptive analyses were conducted for the *colline*, health center, and household data. In the case of the *colline* and household data, the analyses served to provide information on the context of the study area and to ascertain differences among the four study arms at baseline and follow-up (balance among study arms). Double-difference estimates were used to determine the impact of *Tubaramure* on key outcomes at the household level. *Tubaramure* did not specifically aim to improve *colline* characteristics; therefore, no impact of the program at the *colline* level was assessed. In the case of the health centers, which all benefited from *Tubaramure* services (no control group), only baseline and follow-up indicators were compared. All data were analyzed using Stata Release 14.

Colline: Descriptive Analyses

Colline-level results are presented as percentages or means and SDs as appropriate. Results are presented by study arm, and final sample sizes are reported in the results tables in Section 3. To compare results among the four study arms at baseline and follow-up, we used the following linear model[28] for continuous and dichotomous variables, where y_i is the variable or indicator of interest for observation *i,* and three dummy variables (S_i) indicate assignment to one of the three treatment arms:

$$y_i = \beta_0 + \beta_1 S_{i1} + \beta_2 S_{i2} + \beta_3 S_{i3} + \varepsilon_i$$

A joint F-test was used to determine whether there were statistically significant differences among the study arms. For categorical variables, Pearson's chi-square statistic was converted into an F statistic.

Results were considered significantly different among the four study arms if $p < 0.05$. Variables that have significant differences between the study arms are marked with an asterisk (*) in the results tables. For categorical variables, the asterisk is placed in the row of the last category.

Health Center: Descriptive Analyses

Results for the 11 health centers are presented as counts or as means and SDs as appropriate. *Tubaramure* health service interventions were implemented in all health centers regardless of whether they were located in a control or treatment arm; therefore, results are not differentiated by study arm. Because of the small sample size, statistical tests were not used to compare differences between baseline and follow-up.

Household: Descriptive Analyses and Impact Estimation

Descriptive analyses. Similar to the *colline* analyses, the variables or indicators of interest are presented as percentages or means and SDs as appropriate in the household, maternal, and child results sections. All tables present the variables and indicators by study arm, and, when appropriate, results are presented according to child age. The final sample size for each outcome is reported in the tables.

To compare results among the four study arms at baseline and follow-up, we used the following linear model for continuous and dichotomous variables, where y_i is the variable or indicator of interest for observation *i*:

[27] The analyses included child age as a covariate, to account for potential differences in age distribution across study arms.
[28] Note that the study was not powered to detect *colline*-level differences.

$$y_i = \beta_0 + \beta_1 S_{i1} + \beta_2 S_{i2} + \beta_3 S_{i3} + \varepsilon_i$$

We included three dummy variables ($S_{i.}$) for the study arms. For data collected at the household level, the standard errors (SEs) of the parameters were adjusted for the (potential) lack of independence between observations in the same *colline* by using a clustered sandwich estimator. A joint F-test was used to determine whether there were statistically significant differences among the study arms.

For categorical variables, Pearson's chi-square statistic was adjusted for the lack of independence between clusters with the second-order correction of Rao and Scott (1984) and converted into an F statistic.

Results were considered significantly different among the study arms if $p < 0.05$. Variables with significant differences among the study arms are marked with an asterisk (*) in the results tables. For categorical variables, the asterisk is placed in the row of the last category. Differences among the study arms at baseline provided information on the comparability (also referred to as "balance") of the study arms before the start of the program and are commented on in the results sections. Two things must be noted. First, differences between arms at follow-up are also noted in the tables, but not discussed in the text as they might be due to the impact of the program. The impact estimation is discussed in the following section. Second, we did not statistically test the changes between baseline and follow-up within each study arm. Terms in the text, such as "increase" and "decrease," which describe the overall secular trends between baseline and follow-up, are not formally tested as this was not a study objective. We only formally tested the difference in these changes over time between the treatment arms and the control arms as described in the following section.

Impact estimation. Program impact was estimated using the following double-difference[29] *colline*-fixed effect model. This model compares the change in the outcomes from baseline to follow-up among study arms, where T_j is time (baseline or follow-up), S_i is a dummy representing being assigned to any study arm, and C is a vector representing *colline*-level fixed effects.

$$y_{t=0,1} = \beta_0 + \beta_1 T_j + \beta_2 S_i + \beta_3 T_j S_i + \beta_4 C + \beta_i X_{i,t=0} + \varepsilon$$

The coefficient β_3 represents the estimated treatment effect of the program. By using *colline*-level fixed effects,[30] the model controls for unobserved *colline* characteristics that did not change between baseline and follow-up. To reduce residual noise and maximize power, baseline covariates $X_{i,t=0}$ were added to the model for some outcomes. The covariates included in the model are indicated in the footnotes of each table.[31] In addition to the model above, we also estimated the treatment arm-specific impact, by including a dummy variable (and a treatment arm by time interaction term) for each treatment arm. The difference

[29] The double-difference accounts for any differences between the treatment and control groups that are constant over time.

[30] An alternative approach would be to include *colline*-random effects, which is equivalent to adding a *colline*-level error term to the model (in addition to the individual-level error term already present). The disadvantage of the random-effects approach is that *colline*-level effects might be correlated with the independent variables. If this is the case, the assumption that the error term is independent from the independent variables is no longer met, thus creating the problem of endogeneity, which might lead to biased (impact) estimates (Wooldridge 2003).

[31] Covariates were added to the impact models for child Hb, anemia, child anthropometry, and child language and development milestones. These outcomes are well-studied in the literature, and covariates included those commonly used in these models, such as child age, sex, and maternal height for the child anthropometric outcomes. Identifying which covariates to include for all the other outcomes was beyond the scope of this report. Consistent with the first follow-up report, no covariates (other than the *colline*-fixed effects) were included for these other outcomes.

in impact between arms was estimated for child linear growth (HAZ and stunting)[32] by changing the reference group.

As indicated above, data on some outcomes were only available at follow-up (e.g., maternal dietary diversity). The simple-difference model was used to estimate impact on these outcomes:

$$y_{t=1} = \beta_0 + \beta_1 S_i + \beta_i X_{i,t=0} + \varepsilon$$

The coefficient β_1 represents the estimated treatment effect of the program in this model. As for the double-difference model, we also estimated the treatment arm-specific impact by including a dummy for each treatment arm.

The SEs of all estimated parameters were adjusted for the (potential) lack of independence between observations in the same *colline* by using a clustered sandwich estimator. We conducted intent-to-treat analyses. One-sided tests were used when there was a clear *a priori* hypothesis of the direction of the effect; these are indicated in the footnotes of each table.

[32] The study was powered to detect an impact (i.e., a difference between treatment and control) for a number of outcomes (including child linear growth and development); with respect to differences between arms (T24 vs. T18 and T24 vs. TNFP), the study was only powered to detect these for the linear growth outcome.

3. Results: *Colline* and Health Center Characteristics

3.1 *Colline* Characteristics

3.1.1 Utilities, Infrastructure, and Access to Services

At baseline, the four study arms did not differ in their access to utilities. Very few *collines* had access to any form of electricity, and public pumps were the most common source of drinking water during both the dry and rainy seasons (**Table 8**). However, there were some changes in access to utilities and sources of drinking water between 2010 and 2014. Specifically, privately generated electricity (i.e., generators, solar panels, batteries) became more common, such that approximately 40 percent of *collines* had access to some form of electricity by 2014. Likewise, there were shifts in the primary sources of drinking water; the percentage of households that used rivers, lakes, or rainwater as their primary source of drinking water in either the dry or the wet season decreased, and the percentages that used tap water or public pumps as their primary source of drinking water increased. Access to landlines and mobile phones also improved between baseline and follow-up. In 2014, more *collines* reported that the closest landline was within 5 km, when compared with 2010, and by 2014 residents in all *collines* could access a mobile phone network within the *colline*. Additionally, the percentage of *collines* with a place to charge a mobile phone increased from 23.3 percent to 73.6 percent between 2010 and 2014.

At baseline, *colline* leaders reported that more than half of the *collines* were within 5 km of a regular (i.e., more than weekly) market (61.7 percent), a weekly market (70.7 percent), a church (96.7 percent), a bus stop (55.9 percent), or an administrative office (58.3 percent) (**Table 9**). The proximity of these services did not differ across the four study arms at baseline. Access was similar at follow-up, with more than half of the c*ollines* within 5 km of a regular market (59.6 percent), a weekly market (62.1 percent), a church (100 percent), or a bus stop (72.8 percent); nearly half (46.8 percent) were within 5 km of an administrative office.

3.1.2 Transportation

The primary form of transportation within the *collines* and to neighboring *collines* was walking, at both baseline and follow-up (**Table 10**). To travel to other cities and provinces, walking and bicycling were the most common means at baseline. The importance of walking dropped over time, resulting in roughly equal shares for bus/minibus, private car, bicycle, and walking as the most common transportation means at follow-up. At baseline, 70.0 percent of *collines* could access a road within the *colline*; in all but one remaining *colline,* road access was available within 5 km. At follow-up, fewer *collines* (42.3 percent) could access a road within the *colline*, and almost 10 percent could not access a road within 10 km. The vast majority of the closest roads were dirt or laterite at both surveys, and on average they were usable 9 months out of the year. For more than half of the *collines*, the reported distance to the closest asphalt road in both surveys was more than 10 km. None of these transportation characteristics differed significantly among the four study arms at baseline.

3.1.3 Access to Schools and Health Services

At baseline, primary and secondary schools were available in 75.0 percent and 35.0 percent of the collines, respectively; this improved to 81.5 percent and 55.5 percent at follow-up (**Table 11**). The average distance to the nearest primary school for *collines* without one decreased between surveys, from 3.5 km at baseline to 2.5 km at follow-up. For *collines* without secondary schools, the closest one was on average 5 km away at the time of both surveys. Most primary and secondary schools had access to latrines (90.1 percent and 86.6 percent at baseline, and 82.2 percent and 91.6 percent at follow-up, respectively). However, a smaller percentage of schools had access to drinking water—25.3 percent of primary and 50.0 percent of secondary schools at baseline and 35.9 percent of primary and 68.5 percent of secondary schools at follow-up. Access to schools and characteristics of those schools were similar across the four study arms at baseline.

Access to health services was limited, although it improved between 2010 and 2014. In 2010, only 20.0 percent of *collines* had a health center (public or private) (**Table 12**). By 2014, 31.5 percent of *collines* had health centers within their borders, and the percentage of *collines* that had a health center within 5 km increased from 69.0 percent to 75.0 percent. Access to a hospital also improved over time, as the percentage of *collines* with a hospital or within 10 km of one increased from 38.4 percent to 59.5 percent. Most health services were reached on foot at both baseline and follow-up.

Community members were also served by health staff who were not specific to a health center or hospital. At both surveys, most *collines* reported having at least one midwife, and around half of the *collines* were served by three or more midwives. Community health workers served all but one *colline* at baseline and served all of them at follow-up. Access to health services at baseline was similar across the four study arms.

3.1.4 Agriculture

Between 2010 and 2014, the percentage of *collines* growing most of the key crops declined, with the exception of sweet potatoes and beans. At baseline, manioc, maize, sweet potatoes, and beans were grown in nearly all *collines* while the other primary crops were cultivated in 55.0 percent to 83.3 percent of the *collines* (**Table 13**). The most commonly cultivated vegetables in the study *collines* at baseline were amaranth, cabbage, eggplant, peas, tomatoes, and squash. However, as with the cultivation of the key crops, the percentage of *collines* cultivating these vegetables decreased between 2010 and 2014. At baseline, the types of crops cultivated in the study *collines* were similar across study arms.

3.1.5 Social Groups, Development Programs, and Recent Events

The leaders from all *collines* reported the presence of community associations, cooperatives, and other groups. At baseline, 35.0 percent had between one and three organizations, 38.3 percent between four and seven, and 26.7 percent more than eight groups (**Table 15**). A smaller percentage of *collines* had more than eight groups at follow-up, but community groups were still as common as they were in 2010. Nearly all groups counted women among their members, and most had multiple activities. The most common activities were agriculture (83.3 percent at baseline and 70.1 percent at follow-up) and animal husbandry (49.9 percent at baseline and 45.0 percent at follow-up). Less commonly, groups engaged in credit, trade, beekeeping, crafts, and health. The percentage of groups engaged in money lending and health promotion tripled between baseline and follow-up (from 9.4 percent to 27.2 percent and from 2.5 percent to 8.0 percent, respectively). These characteristics did not differ among the four study arms at baseline.

At follow-up, fewer *colline* respondents reported that more people arrived in the *colline* than left over the past 5 years, falling from 76.7 percent at baseline to 66.7 percent (**Table 16**). Regarding weather conditions during the past 12 months, at baseline 51.6 percent of *colline* respondents reported having had more rainfall than usual, whereas at follow-up the majority of *colline* respondents (70.4 percent) reported that there was less rainfall than usual. In addition, the percentage that reported that it was hotter than normal during the previous 12 months increased between 2010 and 2014 (from 75.0 percent to 85.2 percent). Lastly, the percentage of *colline* respondents that reported that living conditions had become worse during the past 3 years dropped from 83.3 percent to 64.8 percent between 2010 and 2014. There were not significant differences between groups at baseline.

The five most common negative events experienced by *collines* in the 2 years preceding each survey were drought, flood, hail, fire, and famine (**Table 17**). The reported occurrence of droughts and hail storms increased between the baseline and follow-up surveys, while the reported occurrence of famine decreased in the T24 and T18 study arms, remained roughly the same in the TNFP study arm, and increased over time in the control arm. There were no significant differences between study arms at baseline.

Table 8. Utilities and infrastructure in each *colline*[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**54**	**41**	**14**	**13**	**14**	**13**
Electricity source												
None	96.7	95.6	93.3	93.3	100.0	100.0	59.3	58.5	50.0	53.8	71.4	61.5
Water company	1.7	2.2	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Individual generator	0.0	0.0	0.0	0.0	0.0	0.0	1.9	2.4	0.0	7.7	0.0	0.0
Solar panels	1.7	2.2	6.7	0.0	0.0	0.0	27.8	29.3	28.6	38.5	21.4	23.1
Batteries	0.0	0.0	0.0	0.0	0.0	0.0	11.1	9.8	21.4	0.0	7.1	15.4
Distance to a landline telephone												
0 km (in *colline*)	10.0	13.3	26.7	6.7	6.7	0.0	14.6	16.7	10.0	22.2	18.2	9.1
1–5 km	11.7	11.1	6.7	6.7	20.0	13.3	12.2	13.3	20.0	0.0	18.2	9.1
6–10 km	5.0	6.7	6.7	13.3	0.0	0.0	9.8	6.7	0.0	11.1	9.1	18.2
> 10 km	73.3	68.9	60.0	73.3	73.3	86.7	63.4	63.3	70.0	66.7	54.5	63.6
Distance to a mobile phone network												
0 km (in *colline*)	88.3	91.1	93.3	93.3	86.7	80.0	100.0	100.0	100.0	100.0	100.0	100.0
1–5 km	6.7	6.7	6.7	0.0	13.3	6.7	0.0	0.0	0.0	0.0	0.0	0.0
6–10 km	3.3	2.2	0.0	6.7	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0
> 10 km	1.7	0.0	0.0	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0
Possible to charge mobile in the *colline*												
	23.3	22.2	20.0	40.0	6.7	26.7	73.6	75.6	78.6	92.3	57.1	66.7
Most common drinking water source: dry season												
Tap water	8.3	6.7	13.3	6.7	0.0	13.3	20.4	19.5	14.3	30.8	14.3	23.1
Public pump	68.3	73.3	73.3	60.0	86.7	53.3	72.2	73.2	85.7	53.8	78.6	69.2
Uncovered well	3.3	2.2	6.7	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0
River or lake	20.0	17.8	6.7	33.3	13.3	26.7	7.4	7.3	0.0	15.4	7.1	7.7
Most common drinking water source: wet season												
Tap water	15.0	13.3	20.0	20.0	0.0	20.0	22.2	19.5	14.3	30.8	14.3	30.8
Public pump	66.7	73.3	66.7	66.7	86.7	46.7	70.4	73.2	85.7	53.8	78.6	61.5
Uncovered well	3.3	2.2	6.7	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0
Rainwater	6.7	2.2	0.0	6.7	0.0	20.0	5.6	4.9	0.0	7.7	7.1	7.7
River or lake	8.3	8.9	6.7	6.7	13.3	6.7	1.9	2.4	0.0	7.7	0.0	0.0

[a] Values are %.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up ranged from N = 41 to 54 in the full sample; N = 30 to 41 in the treatment arms; N = 10 to 14 in the T24 arm; N = 9 to 13 in the T18 arm; N = 11 to 14 in the TNFP arm; and N = 11 to 13 in the control arm.
* There were no statistically significant differences among study arms, $p < 0.05$.

Table 9. Distance to services[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**52**	**39**	**14**	**13**	**14**	**13**
Distance to closest regular market												
0 km (in *colline*)	11.7	13.3	20.0	6.7	13.3	6.7	23.1	23.1	15.4	23.1	30.8	23.1
1–5 km	50.0	48.9	46.7	40.0	60.0	53.3	36.5	35.9	53.8	23.1	30.8	38.5
6–10 km	18.3	17.8	20.0	20.0	13.3	20.0	25.0	25.6	23.1	23.1	30.8	23.1
> 10 km	20.0	20.0	13.3	33.3	13.3	20.0	15.4	15.4	7.7	30.8	7.7	15.4
Distance to closest weekly market												
0 km (in *colline*)	34.5	25.6	23.1	40.0	13.3	60.0	29.7	26.9	25.0	44.4	11.1	36.4
1–5 km	36.2	39.5	38.5	40.0	40.0	26.7	32.4	26.9	37.5	22.2	22.2	45.5
6–10 km	15.5	16.3	15.4	13.3	20.0	13.3	32.4	42.3	25.0	33.3	66.7	9.1
> 10 km	13.8	18.6	23.1	6.7	26.7	0.0	5.4	3.8	12.5	0.0	0.0	9.1
Distance to closest church												
0 km (in *colline*)	85.0	84.4	93.3	93.3	66.7	86.7	92.2	92.1	92.9	100.0	85.7	92.3
1–5 km	11.7	11.1	6.7	6.7	20.0	13.3	7.8	7.9	7.1	0.0	14.3	7.7
6–10 km	1.7	2.2	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
> 10 km	1.7	2.2	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Distance to closest bus stop												
0 km (in *colline*)	22.0	22.7	26.7	33.3	7.1	20.0	45.5	46.9	50.0	70.0	25.0	41.7
1–5 km	33.9	34.1	26.7	26.7	50.0	33.3	27.3	25.0	40.0	0.0	33.3	33.3
6–10 km	11.9	13.6	13.3	13.3	14.3	6.7	20.5	25.0	10.0	20.0	41.7	8.3
> 10 km	32.2	29.5	33.3	26.7	28.6	40.0	6.8	3.1	0.0	10.0	0.0	16.7
Distance to closest administrative center												
0 km (in *colline*)	13.3	17.8	26.7	20.0	6.7	0.0	17.0	20.6	18.2	30.0	15.4	7.7
1–5 km	45.0	44.4	26.7	46.7	60.0	46.7	29.8	29.4	45.5	10.0	30.8	30.8
6–10 km	20.0	22.2	46.7	6.7	13.3	13.3	36.2	38.2	36.4	30.0	46.2	30.8
> 10 km	21.7	15.6	0.0	26.7	20.0	40.0	17.0	11.8	0.0	30.0	7.7	30.8

[a] Values are %.
[b] Sample size at baseline ranged from N = 58 to 60 in the full sample; N = 43 to 45 in the treatment arms; N = 13 to 15 in the T24 arm; and N = 14 to 15 in the TNFP arm.
[c] Sample size at follow-up ranged from N = 37 to 52 in the full sample; N = 26 to 39 in the treatment arms; N = 8 to 14 in the T24 arm; N = 9 to 13 in the T18 arm; N = 9 to 14 in the TNFP arm; and N = 11 to 13 in the control arm.
* There were no statistically significant differences among study arms, $p < 0.05$.

Table 10. Local forms of transportation[a]

	Baseline[b]						Follow-up[c]					
	Full sample	Study arm					Full sample	Study arm				
		All treated	T24	T18	TNFP	Control		All treated	T24	T18	TNFP	Control
N	60	45	15	15	15	15	54	41	14	13	14	13
Primary form of transportation within the *colline*												
Bus or minibus	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.7	0.0	0.0	7.7	0.0
Motorcycle	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bicycle	3.3	4.4	13.3	0.0	0.0	0.0	4.0	2.7	8.3	0.0	0.0	7.7
Walk	96.7	95.6	86.7	100.0	100.0	100.0	94.0	94.6	91.7	100.0	92.3	92.3
Primary form of transportation to nearby *collines*												
Bus or minibus	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.8	0.0	0.0	8.3	0.0
Private car	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Motorcycle	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bicycle	3.3	4.4	6.7	0.0	6.7	0.0	14.3	13.9	16.7	16.7	8.3	15.4
Walk	96.7	95.6	93.3	100.0	93.3	100.0	83.7	83.3	83.3	83.3	83.3	84.6
Primary form of transportation to other cities and provinces												
Bus or minibus	15.0	15.6	33.3	6.7	6.7	13.3	20.0	15.8	25.0	7.7	15.4	33.3
Taxi	1.7	0.0	0.0	0.0	0.0	6.7	2.0	2.6	8.3	0.0	0.0	0.0
Private car	10.0	6.7	0.0	13.3	6.7	20.0	24.0	21.1	16.7	23.1	23.1	33.3
Motorcycle	5.0	6.7	0.0	6.7	13.3	0.0	10.0	13.2	0.0	23.1	15.4	0.0
Bicycle	21.7	22.2	20.0	20.0	26.7	20.0	20.0	21.1	25.0	30.8	7.7	16.7
Walk	46.7	48.9	46.7	53.3	46.7	40.0	24.0	26.3	25.0	15.4	38.5	16.7
Distance to closest road												
0 km (in *colline*)	70.0	68.9	73.3	73.3	60.0	73.3	42.3	46.2	46.2	33.3	57.1	30.8
1–5 km	28.3	28.9	26.7	26.7	33.3	26.7	48.1	46.2	46.2	50.0	42.9	53.8
6–10 km	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
> 10 km	1.7	2.2	0.0	0.0	6.7	0.0	9.6	7.7	7.7	16.7	0.0	15.4
Material of closest road												
Asphalt	5.0	6.7	0.0	13.3	6.7	0.0	11.1	12.2	7.1	15.4	14.3	7.7
Dirt	86.7	86.7	93.3	80.0	86.7	86.7	63.0	68.3	78.6	53.8	71.4	46.2
Laterite	8.3	6.7	6.7	6.7	6.7	13.3	24.1	17.1	14.3	30.8	7.1	46.2
Paved with stone	0.0	0.0	0.0	0.0	0.0	0.0	1.9	2.4	0.0	0.0	7.1	0.0
Number of months vehicles could be used on closest road last year												
Simple car	8.3 ± 4.7	8.7 ± 4.4	9.7 ± 3.5	9.5 ± 4.3	7.0 ± 5.2	7.1 ± 5.2	9.2 ± 4.3	8.9 ± 4.6	8.0 ± 5.4	9.4 ± 4.6	9.2 ± 4.2	10.3 ± 2.8
Four-wheel drive vehicle	9.9 ± 3.8	10.1 ± 3.6	10.9 ± 2.3	10.7 ± 3.0	8.7 ± 5.0	9.2 ± 4.5	9.0 ± 4.4	9.0 ± 4.5	8.1 ± 5.2	8.5 ± 5.0	10.2 ± 3.2	9.1 ± 4.4
Bus	6.8 ± 5.3	7.2 ± 5.2	6.8 ± 5.2	8.5 ± 4.9	6.4 ± 5.7	5.5 ± 5.6	8.4 ± 4.9	7.8 ± 5.3	7.8 ± 5.7	7.5 ± 5.6	8.0 ± 5.0	10.3 ± 2.8
Truck	8.0 ± 4.7	8.0 ± 4.7	8.7 ± 4.8	8.5 ± 4.3	6.8 ± 5.1	8.1 ± 4.8	7.8 ± 5.2	7.4 ± 5.4	6.5 ± 6.1	7.4 ± 5.4	8.1 ± 5.1	9.0 ± 4.0
Distance to closest asphalt road												
0 km (in *colline*)	6.9	7.0	6.7	13.3	0.0	6.7	13.3	17.6	16.7	20.0	16.7	0.0
1–5 km	20.7	18.6	26.7	20.0	7.7	26.7	17.8	11.8	25.0	0.0	8.3	36.4
6–10 km	3.4	4.7	0.0	0.0	15.4	0.0	15.6	17.6	16.7	0.0	33.3	9.1
> 10 km	69.0	69.8	66.7	66.7	76.9	66.7	53.3	52.9	41.7	80.0	41.7	54.5

[a] Values are % or mean ± SD.
[b] Sample size at baseline ranged from N = 58 to 60 in the full sample; N = 43 to 45 in the treatment arms; and N = 13 to 15 in the TNFP arm.

[c] Sample size at follow-up ranged from N = 45 to 54 in the full sample; N = 34 to 41 in the treatment arms; N = 12 to 14 in the T24 arm; N = 10 to 13 in the T18 arm; N = 12 to 14 in the TNFP arm; and N = 11 to 13 in the control arm.
* There were no statistically significant differences among study arms, p < 0.05.

Table 11. School characteristics[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**54**	**41**	**14**	**13**	**14**	**13**
% of *collines* with … primary schools												
None	25.0	26.7	33.3	6.7	40.0	20.0	18.5	22.0	14.3	30.8	21.4	7.7
One	51.7	46.7	60.0	46.7	33.3	66.7	57.4	53.7	71.4	30.8	57.1	69.2
Two or more	23.3	26.7	6.7	46.7	26.7	13.3	24.1	24.4	14.3	38.5	21.4	23.1
Primary school characteristics												
Has a latrine	90.1	90.5	94.4	93.3	83.3	88.9	86.6	87.9	98.6	86.4	77.7	82.4
Has drinking water	25.3	25.4	36.0	17.9	22.0	25.0	35.9	33.2	33.6	31.8	33.8	44.6
% of *collines* with … secondary schools												
None	65.0	60.0	53.3	60.0	66.7	80.0	44.4	48.8	57.1	38.5	50.0	30.8
One	28.3	31.1	40.0	26.7	26.7	20.0	40.7	34.1	35.7	30.8	35.7	61.5
Two or more	6.7	8.9	6.7	13.3	6.7	0.0	14.8	17.1	7.1	30.8	14.3	7.7
Secondary school characteristics												
Has a latrine	82.2	84.1	86.1	73.8	92.4	76.7	91.6	92.0	96.4	92.3	87.4	90.4
Has drinking water	50.0	55.6	66.7	50.9	49.1	34.4	68.5	67.9	74.6	68.5	60.7	70.3
Distance to primary school if none in the *colline* (km)												
	3.5 ± 1.8	3.3 ± 1.6	3.0 ± 1.5	3.3 ± 1.8	3.5 ± 1.7	4.1 ± 2.3	2.5 ± 1.0	2.5 ± 1.0	2.0 ± 1.4	2.5 ± 0.7	3.0 ± 1.4	–
Distance to secondary school if none in the *colline* (km)												
	4.8 ± 2.3	4.9 ± 2.5	3.9 ± 1.9	5.4 ± 2.5	5.4 ± 3.0	4.4 ± 1.7	4.7 ± 2.8	4.8 ± 3.0	4.0 ± 1.7	6.8 ± 5.6	4.7 ± 2.1	4.3 ± 1.3

[a] Values are % or mean ± SD.
[b] Sample size at baseline ranged from N = 37 to 60 in the full sample; N = 25 to 45 in the treatment arms; N = 8 to 15 in the T24 arm; N = 8 to 15 in the T18 arm; N = 9 to 15 in the TNFP arm; and N = 12 to 15 in the control arm.
[c] Sample size at follow-up ranged from N = 6 to 54 in the full sample; N = 6 to 41 in the treatment arms; N = 2 to 14 in the T24 arm; N = 2 to 13 in the T18 arm; N = 2 to 14 in the TNFP arm; and N = 0 to 13 in the control arm.
* There were no statistically significant differences among study arms, p < 0.05.

Table 12. Access to health services by *colline* residents[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	60	45	15	15	15	15	54	41	14	13	14	13
Facilities available in the *colline*												
Public health center	16.7	17.8	13.3	20.0	20.0	13.3	28.3	24.4	14.3	38.5	21.4	41.7
Any health center (public or private)	20.0	22.2	13.3	26.7	26.7	13.3	31.5	29.3	14.3	46.2	28.6	38.5
Hospital (all public)	3.3	2.2	0.0	0.0	6.7	6.7	1.9	2.4	0.0	0.0	7.1	0.0
Distance to closest health center												
0 km (in *colline)*	20.7	23.3	14.3	26.7	28.6	13.3	34.6	32.5	23.1	46.2	28.6	41.7
1–5 km	48.3	46.5	64.3	40.0	35.7	53.3	40.4	40.0	53.8	38.5	28.6	41.7
6–10 km	20.7	20.9	14.3	20.0	28.6	20.0	21.2	22.5	23.1	7.7	35.7	16.7
> 10 km	10.3	9.3	7.1	13.3	7.1	13.3	3.8	5.0	0.0	7.7	7.1	0.0
Distance to closest hospital												
0 km (in *colline*)	3.8	2.4	0.0	0.0	7.7	9.1	6.3	8.3	0.0	0.0	22.2	0.0
1–5 km	19.2	19.5	30.8	6.7	23.1	18.2	18.8	12.5	22.2	0.0	11.1	37.5
6–10 km	15.4	12.2	7.7	13.3	15.4	27.3	34.4	41.7	44.4	16.7	55.6	12.5
> 10 km	61.5	65.9	61.5	80.0	53.8	45.5	40.6	37.5	33.3	83.3	11.1	50.0
Transportation to health centers												
Walking	97.5	97.4	92.9	100.0	100.0	97.8	91.5	88.9	91.7	79.5	96.4	100.0
Bicycle	2.5	2.6	7.1	0.0	0.0	2.2	5.4	7.0	8.3	9.0	3.6	0.0
Transportation to hospitals												
Walking	69.3	64.6	67.9	64.4	61.1	86.4	73.7	66.7	70.4	50.0	75.0	93.8
Bicycle	16.9	17.8	26.7	11.1	16.7	13.6	16.7	22.5	18.5	36.1	16.7	0.0
Private car	12.8	16.3	1.5	24.4	22.2	0.0	6.5	8.7	11.1	5.6	8.3	0.0
% of *collines* with … midwives												
None	10.3	7.0	6.7	0.0	14.3	20.0	13.7	15.8	8.3	7.7	30.8	7.7
One	15.5	16.3	20.0	14.3	14.3	13.3	7.8	7.9	0.0	15.4	7.7	7.7
Two	24.1	23.3	13.3	28.6	28.6	26.7	19.6	21.1	25.0	7.7	30.8	15.4
Three or more	50.0	53.5	60.0	57.1	42.9	40.0	58.8	55.3	66.7	69.2	30.8	69.2
% of *collines* with … community health workers												
None	1.7	2.2	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
One	1.7	2.2	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Two	70.0	64.4	60.0	60.0	73.3	86.7	11.1	12.2	14.3	7.7	14.3	7.7
Three or more	26.7	31.1	40.0	40.0	13.3	13.3	88.9	87.8	85.7	92.3	85.7	92.3

[a] Values are %.
[b] Sample size at baseline ranged from N = 51 to 60 in the full sample; N = 40 to 45 in the treatment arms; N = 13 to 15 in the T24 arm; N = 13 to 15 in the T18 arm; N = 12 to 15 in the TNFP arm; and N = 11 to 15 in the control arm.
[c] Sample size at follow-up ranged from N = 31 to 54 in the full sample; N = 23 to 41 in the treatment arms; N = 9 to 14 in the T24 arm; N = 6 to 13 in the T18 arm; N = 8 to 14 in the TNFP arm; and N = 8 to 13 in the control arm.
* There were no statistically significant differences among study arms, $p < 0.05$.

Table 13. Most common crops and vegetables[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**54**	**41**	**14**	**13**	**14**	**13**
Most common crops												
Manioc	100.0	100.0	100.0	100.0	100.0	100.0	92.6	90.2	85.7	84.6	100.0	100.0
Maize	95.0	93.3	93.3	86.7	100.0	100.0	83.3	78.0	71.4	69.2	92.9	100.0
Sweet potatoes	91.7	91.1	93.3	86.7	93.3	93.3	94.4	97.6	100.0	100.0	92.9	84.6
Beans	90.0	88.9	86.7	93.3	86.7	93.3	96.3	95.1	100.0	84.6	100.0	100.0
Peanuts	83.3	80.0	73.3	80.0	86.7	93.3	70.4	70.7	50.0	76.9	85.7	69.2
Sorghum	78.3	77.8	60.0	80.0	93.3	80.0	57.4	56.1	42.9	61.5	64.3	61.5
Rice	61.7	55.6	53.3	60.0	53.3	80.0	59.3	56.1	50.0	76.9	42.9	69.2
Pigeon peas	58.3	60.0	60.0	60.0	60.0	53.3	51.9	53.7	35.7	69.2	57.1	46.2
Eleusine	56.7	57.8	60.0	60.0	53.3	53.3	40.7	46.3	57.1	30.8	50.0	23.1
Potatoes	55.0	60.0	73.3	53.3	53.3	40.0	40.7*	48.8	71.4	38.5	35.7	15.4
Most common vegetables												
Amaranth (*lengalenga*)	53.3	51.1	60.0	33.3	60.0	60.0	50.0	53.7	71.4	53.8	35.7	38.5
Cabbage	45.0	51.1	66.7	33.3	53.3	26.7	25.9	26.8	50.0	15.4	14.3	23.1
Eggplant	38.3	40.0	46.7	33.3	40.0	33.3	24.1	24.4	28.6	30.8	14.3	23.1
Peas	38.3	40.0	60.0	33.3	26.7	33.3	14.8	17.1	7.1	15.4	28.6	7.7
Tomatoes	35.0	35.6	46.7	20.0	40.0	33.3	20.4*	22.0	42.9	23.1	0.0	15.4
Squash	21.7	17.8	26.7	6.7	20.0	33.3	9.3	9.8	21.4	7.7	0.0	7.7

[a] Values are %.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
* Study arms differed, $p < 0.05$.

Table 14. Ten most common fruit trees[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**54**	**41**	**14**	**13**	**14**	**13**
Avocado	98.3	97.8	100.0	93.3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Mango	85.0	82.2	80.0	80.0	86.7	93.3	87.0	90.2	92.9	84.6	92.9	76.9
Banana[d]	80.0	77.8	86.7	73.3	73.3	86.7	-	-	-	-	-	-
Pineapple	75.0	75.6	80.0	73.3	73.3	73.3	66.7	70.7	71.4	61.5	78.6	53.8
Orange	33.3	35.6	40.0	40.0	26.7	26.7	31.5	31.7	14.3	38.5	42.9	30.8
Guava	28.3	31.1	33.3	33.3	26.7	20.0	22.2	24.4	28.6	30.8	14.3	15.4
Japanese plum	26.7	26.7	33.3	13.3	33.3	26.7	20.4	22.0	35.7	7.7	21.4	15.4
Lemon	25.0	22.2	33.3	20.0	13.3	33.3	25.9	26.8	14.3	46.2	21.4	23.1
Coffee[d]	23.3	22.2	26.7	6.7	33.3	26.7	-	-	-	-	-	-
Papaya	25.0	22.2	26.7	20.0	20.0	33.3	27.8	24.4	7.1	38.5	28.6	38.5

[a] Values are %.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
[d] Not reported at follow-up due to data errors.
* There were no statistically significant differences among study arms, $p < 0.05$.

Table 15. Presence of associations, cooperatives, and other groups[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**54**	**41**	**14**	**13**	**14**	**13**
Community groups												
1–3 groups	35.0	37.8	40.0	40.0	33.3	26.7	48.1	39.0	50.0	23.1	42.9	76.9
4–7 groups	38.3	37.8	26.7	40.0	46.7	40.0	46.3	56.1	50.0	76.9	42.9	15.4
8+ groups	26.7	24.4	33.3	20.0	20.0	33.3	5.6	4.9	0.0	0.0	14.3	7.7
Have women members	97.6	97.0	97.8	96.9	96.4	99.0	99.5	99.4	100.0	98.1	100.0	100.0
Primary activities												
Agriculture	83.3	86.9	93.6	82.1	84.5	73.2	70.1	68.9	75.0	70.8	61.0	73.8
Animal husbandry	49.9	52.3	52.3	47.9	56.8	43.0	45.0	41.7	35.2	49.4	41.1	55.2
Money lending	9.4	6.7	10.3	4.5	5.0	17.3	27.2	28.8	24.4	23.6	38.2	22.1
Trade	5.9	6.9	10.1	3.2	7.1	3.0	4.2	5.6	6.0	9.7	1.2	0.0
Bee keeping	6.2	4.0	4.2	6.0	1.8	12.4	2.3	3.0	5.2	3.8	0.0	0.0
Crafts	2.6	1.9	0.8	1.4	3.4	4.8	1.3	1.3	1.2	0.0	2.5	1.5
Health promotion	2.5	2.8	3.3	1.4	3.6	1.7	8.0	10.5	8.8	7.3	15.1	0.0

[a] Values are %.
[b] Sample size at baseline ranged from N = 58 to 60 in the full sample; N = 43 to 45 in the treatment arms; N = 14 to 15 in the T18 arm; and N = 14 to 15 in the TNFP arm.
[c] Sample size at follow-up did not vary.
* There were no statistically significant differences among study arms, $p < 0.05$.

Table 16. Recent historical events experienced by the *collines*[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**54**	**41**	**14**	**13**	**14**	**13**
Migration during past 5 years												
More arrivals	76.7	75.6	73.3	73.3	80.0	80.0	66.7	58.5	64.3	61.5	50.0	92.3
More departures	11.7	11.1	13.3	13.3	6.7	13.3	18.5	22.0	14.3	38.5	14.3	7.7
Similar arrivals and departures	3.3	4.4	6.7	6.7	0.0	0.0	3.7	4.9	0.0	0.0	14.3	0.0
Neither arrivals nor departures	8.3	8.9	6.7	6.7	13.3	6.7	11.1	14.6	21.4	0.0	21.4	0.0
Rainfall during past 12 months												
A lot more than usual	48.3	55.6	73.3	46.7	46.7	26.7	22.2	19.5	28.6	7.7	21.4	30.8
A little more than usual	3.3	2.2	0.0	6.7	0.0	6.7	7.4	9.8	7.1	0.0	21.4	0.0
Same as usual	5.0	6.7	0.0	6.7	13.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A little less than usual	15.0	8.9	13.3	0.0	13.3	33.3	31.5	31.7	28.6	53.8	14.3	30.8
A lot less than usual	28.3	26.7	13.3	40.0	26.7	33.3	38.9	39.0	35.7	38.5	42.9	38.5
Temperature during past 12 months												
A lot hotter than usual	63.3	57.8	53.3	73.3	46.7	80.0	51.9	48.8	42.9	69.2	35.7	61.5
A little hotter than usual	11.7	11.1	13.3	13.3	6.7	13.3	33.3	39.0	42.9	23.1	50.0	15.4
Same as usual	20.0	24.4	26.7	0.0	46.7	6.7	5.6	0.0	0.0	0.0	0.0	23.1
A little cooler than usual	3.3	4.4	6.7	6.7	0.0	0.0	9.3	12.2	14.3	7.7	14.3	0.0
A lot cooler than usual	1.7	2.2	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Living conditions during past 5 years												
Improved	13.3	17.8	13.3	13.3	26.7	0.0	22.2	22.0	14.3	38.5	14.3	23.1
Worse	83.3	80.0	80.0	86.7	73.3	93.3	64.8	63.4	57.1	53.8	78.6	69.2
Similar	3.3	2.2	6.7	0.0	0.0	6.7	13.0	14.6	28.6	7.7	7.1	7.7

[a] Values are %.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
* There were no statistically significant differences among study arms, $p < 0.05$.

Table 17. Most common negative events experienced by *collines* during preceding 2 years[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**60**	**45**	**15**	**15**	**15**	**15**	**52**	**40**	**14**	**12**	**14**	**12**
Drought	33.3	37.8	46.7	26.7	40.0	20.0	57.7	62.5	64.3	58.3	64.3	41.7
Flood	30.0	35.6	26.7	33.3	46.7	13.3	17.3	17.5	21.4	8.3	21.4	16.7
Hail	21.7	26.7	40.0	26.7	13.3	6.7	38.5	35.0	42.9	41.7	21.4	50.0
Fire	21.7	20.0	33.3	20.0	6.7	26.7	7.7	10.0	0.0	16.7	14.3	0.0
Famine	15.0	13.3	20.0	13.3	6.7	20.0	9.6*	2.5	0.0	0.0	7.1	33.3

[a] Values are %.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
* Study arms differed, $p < 0.05$.

3.2 Health Center Characteristics[33]

3.2.1 Infrastructure, Personnel, and Services

Changes from baseline to follow-up in the basic infrastructure of the 11 health centers—which provide health promotion as well as preventive and lower-level curative care services—were mostly small. Ten centers had a covered waiting area at both surveys (**Table 18**). The number of health centers able to keep patients overnight for observation increased from eight to nine between baseline and follow-up, and the number with consultation rooms specifically for children under 5 years increased from one to seven. At follow-up, all health centers had electricity:[34] six relied on solar panels alone, one relied on an electric grid, and four had access to both sources of electricity. In the event of power outages, four health centers had backup generators.

The availability of trained staff increased between 2010 and 2014. In 2010, only five health centers staffed an A2-level nurse (4 years of nurse training after 10th grade). This doubled by 2014, with 10 centers having an A2-level nurse. All health centers had an A3-level nurse (2 years of nurse training after 10th grade) in both 2010 and 2014; however, only one (in 2014) was specifically trained in nutrition. The number of nurse's aides remained the same (nine in 2010 and 2014). At baseline, no A2-level laboratory technicians (4 years of training after 10th grade) were available, and seven centers employed assistant laboratory technicians. A2-level laboratory technicians were available at three health centers at follow-up, and the number with an assistant lab technician remained the same. The availability of health promoters and community health workers increased only slightly between baseline and follow-up (from 8 to 9 and from 8 to 11, respectively).

Table 18. Health center infrastructure and personnel[a]

N	Baseline[b] 11	Follow-up[c] 11
Infrastructure		
Covered waiting area	10	10
Will keep patients overnight for observations	8	9
Consultation room for children under 5 years	1	7
Electricity[d]		
Only solar panels	–	6
Only electricity from a grid	–	1
Both solar panels and grid	–	4
No source of power	–	0
Generator	–	4
Health care personnel (≥ 1)		
A2-level nurse (4 years of nurse training)	5	10
A3-level nurse (2 years of nurse training)	11	11
Nutritionist (specialized A3-level nurse)	0	1
Nurse's aide	9	9
A2-level laboratory technician (4 years of training)	0	3
Assistant laboratory technician	7	7
Health promoter	8	9
Number of community health workers		
0	3	0
1–5	2	2
6 or more	6	9

[a] Values are counts (i.e., number of health centers). No statistical tests were conducted.

[33] Note that not all changes to health services were attributable to the project. Changes between baseline and follow-up may reflect investments by the government or other agencies. See page 4 for a description of support provided by the project to the health centers.

[34] This was not measured at baseline.

[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
[d] Information not collected at baseline.

Nearly all 11 health centers offered growth monitoring, immunizations, and vitamin A supplements at baseline (**Table 19**). These three services were all offered on average 2 to 3 days a week. Two health centers began offering mobile growth monitoring services (i.e., facility staff traveled to another location to provide services) by follow-up, one of which was doing so in place of the in-facility growth monitoring services it previously offered. Mobile growth monitoring services were infrequently offered, on average less than once a month. Routine deworming[35] was available at 10 health centers at baseline and 8 at follow-up, on average about 1.5 days a week. Curative services for sick children were offered at 11 health centers at baseline and 10 at follow-up, and provided approximately 7 days a week at both surveys.

Pregnant women could access prenatal services at all 11 health centers (approximately 3 days a week), and 9 health centers provided delivery assistance in the health center (around 6 days a week) at both surveys. The number of health centers offering postpartum care increased from 10 to 11, but there was a decline in the average frequency of these services (from 5.4 to 4 days a week).

Table 19. Availability and frequency of services[a]

N	Baseline[b] 11	Follow-up[c] 11
Preventive services for children under 5 years		
Growth monitoring available	11	10
Growth monitoring (days per week)	2.3 ± 2.4	3.1 ± 2.7
Mobile growth monitoring available[d]	–	2
Mobile growth monitoring (days per week)[d]	–	0.1 ± 0.2
Immunizations available	11	11
Immunizations (days per week)	2.2 ± 0.7	2.0 ± 0.6
Vitamin A supplementation available	11	8
Vitamin A supplementation (days per week)	2.2 ± 0.8	1.1 ± 1.1
Routine deworming available	10	8
Routine deworming (days per week)	1.5 ± 1.1	1.7 ± 2.8
Curative services for children under 5 years		
Treatment of sick children	11	10
Treatment of sick children (days per week)	6.5 ± 1.8	7.0 ± 0.0
Maternal health services		
Prenatal care available	11	11
Prenatal care (days per week)	3.1 ± 1.3	3.0 ± 1.7
Delivery assistance available	9	9
Delivery assistance (days per week)	6.3 ± 2.0	5.7 ± 2.8
Postpartum care available	10	11
Postpartum care (days per week)	5.4 ± 2.3	4.0 ± 2.9

[a] Values are counts (i.e., number of health centers) or mean ± SD. No statistical tests were conducted.
[b] Sample size at baseline ranged from N = 9 to 11; mean days per week was calculated only for health centers providing that service.
[c] Sample size at follow-up ranged from N = 10 to 11; mean days per week was calculated only for health centers providing that service.
[d] Information not collected at baseline.

3.2.2 Components of Services

The number of centers offering specific services for sick children improved considerably between baseline and follow-up for all of the services that were not already universally available at baseline (**Table 20**). The mean number of services offered improved from 4.7 at baseline to 7.1 at follow-up. Importantly, health education and the evaluation of immunization and vitamin A supplementation status

[35] Deworming is recommended starting at 1 year of age.

all improved. The availability of four specific services for children with diarrhea also improved substantially. The mean number of services available for children with diarrhea increased from 1.1 at baseline to 3.5 at follow-up.

Table 20. Components of services for sick children[a]

N	Baseline[b] 11	Follow-up[c] 11
Services available to sick children		
Measure body temperature	11	11
Measure weight	10	11
Offer antibiotics	9	11
Offer acetaminophen or wet sponge for fever	6	11
Chart weight	5	7
Provide health education	4	11
Evaluate immunization status	4	10
Evaluate vitamin A status	3	6
Total number of services for sick children (range 0 to 8)	4.7 ± 1.0	7.1 ± 1.0
Services available to children with diarrhea		
Oral rehydration therapy at health center	8	10
ORS packets	2	11
Onsite intravenous solution if severe diarrhea	1	9
Zinc supplementation	1	9
Total number of services for diarrhea (range 0 to 4)	1.1 ± 0.9	3.5 ± 0.7

[a] Values are counts (i.e., number of health centers) or mean ± SD. No statistical tests were conducted.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.

For women seeking prenatal care there were small improvements in the availability of services (**Table 21**). The mean number of services offered improved from 5.9 at baseline to 7.1 at follow-up. This improvement was primarily attributable to the delivery of health education sessions and the administration of deworming treatment.

Table 21. Components of services for pregnant and lactating women[a]

N	Baseline[b] 11	Follow-up[c] 11
Prenatal care components		
Measure weight	11	11
Measure abdomen	10	11
Measure blood pressure	10	11
Offer tetanus vaccine	10	11
Distribute iron folate or iron and folate	10	11
Offer education sessions	8	11
Administer albendazole or mebendazole for deworming	6	11
Administer niclosamide for *taenia* deworming	0	1
Total number of prenatal care components (range 0 to 8)	5.9 ± 0.9	7.1 ± 0.3

[a] Values are counts (i.e., number of health centers) or mean ± SD. No statistical tests were conducted.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.

With regard to laboratory services (**Table 22**), the number of health centers offering HIV tests to pregnant mothers increased from 5 to 11 between baseline and follow-up. Urine and anemia tests were offered for pregnant women at few health centers: urine tests were offered at only three health centers at both baseline and follow-up; the availability of anemia tests increased from two to five health centers. The number of health centers offering blood tests and stool tests to children under 5 increased between baseline and follow-up—from 8 to 10 and from 4 to 10, respectively. Only one health center offered urine tests for children under 5 at the time of both surveys.

Table 22. Laboratory services[a]

N	Baseline[b] 11	Follow-up[c] 11
Prenatal laboratory services		
HIV test	5	11
Urine test	3	3
Anemia test	2	5
Laboratory services for children under 5 years		
Blood test	8	10
Stool test	4	10
Urine test	1	1

[a] Values are counts (i.e., number of health centers). No statistical tests were conducted.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up ranged from N = 10 to 11.

The number of health centers offering blood tests to confirm malaria increased from 7 to 11 between baseline and follow-up (**Table 23**). Bed nets were distributed to pregnant women and to children under 9 months at (nearly) all health centers at the time of both surveys. Bed nets were not distributed to children 9 months and older at either baseline or follow-up.

Table 23. Malaria-related services[a]

N	Baseline[b] 11	Follow-up[c] 11
Blood test for malaria	7	11
Bed net distribution		
Infants (under 9 months)	11	11
Pregnant women at prenatal visits	10	11
Children 9 months and older	0	0

[a] Values are counts (i.e., number of health centers). No statistical tests were conducted.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.

3.2.3 Availability of Equipment, Medication, and Supplies

At baseline, all 11 health centers reported maintaining registries and having immunization cards available (**Table 24**). When further specified at follow-up, all 11 health centers maintained registries of curative consultations, growth monitoring, vaccinations, and prenatal care. However, blank maternal health cards, child health cards, and child health passports were not universally available.

Salter scales and measuring boards were almost universally available at both surveys. Infant scales and mid-upper arm circumference tapes were available at 9 centers at baseline and 10 and 11, respectively, at follow-up. A weight-for-height growth chart was present at only one health center at baseline and three at follow-up.

At both surveys, all health centers had icepacks, thermos carriers, and temperature charts for the refrigerator, which are all required to maintain vaccinations at safe temperatures. The two centers without a working refrigerator at baseline had one in working condition at follow-up.

The availability of equipment for obstetric care remained constant or improved. In both 2010 and 2014, a *Pinard* obstetric stethoscope, adult blood pressure cuff, and measuring tape were universally available, and an adult scale was present at 10 health centers, a delivery table at 9, and a pelvimeter at 7. The availability of vaginal speculums increased from 10 to 11 centers, the number with a gynecological table from 7 to 10, and the number with a gynecological lamp from 2 to 3. A hemoglobin tester, which was previously not available at any health center, was present in five health centers at follow-up. Also at follow-up, six health centers had wheelchairs.

The health centers were well-equipped with sterile equipment and materials. Gloves, disinfectant, sutures, disposable needles and syringes, and containers for disposing of medical waste were universally available at both surveys. Sterile dressings, a sink, and soap (assessed only at follow-up) were available in all health centers, and trash cans were available at 10. The availability of wash drums improved between the surveys (from 10 to 11 centers), and 7 centers had wash bottles at both surveys. Cold decontamination containers were available at fie health centers at baseline and eight at follow-up. At follow-up, only seven health centers had sterile towels/napkins available.

For the diagnostic and curative care of children, stethoscopes and thermometers were universally available at both surveys. The number with otoscopes and examination tables increased from 10 to 11. A child blood pressure cuff was available at only three health centers at baseline and one at follow-up, and a laryngeal mirror was available at one health center at baseline and three at follow-up. At follow-up, nine health centers had a flashlight, nine had a suction bulb or electric suction device, and eight had a metal tongue depressor.

Table 24. Availability of equipment[a]

N	Baseline[b] 11	Follow-up[c] 11
Waiting room and office furniture		
Bench	11	10
Stool	3	11
Chairs[e]	–	11
Office table[e]	–	11
Record keeping		
Registries[d]	11	–
Registry of curative consultation[e]	–	11
Registry of growth monitoring[e]	–	11
Registry of vaccinations[e]	–	11
Registry of prenatal care[e]	–	11
Immunization card[d]	11	–
Maternal health card[e]	–	7
Child health card[e]	–	9
Child health passport[e]	–	8
Growth monitoring equipment		
Salter scale	11	11
Measuring board	11	10
Infant scale	9	10
Mid-upper arm circumference tape	9	11
Weight-for-height growth chart	1	3
Immunization equipment		
Ice pack	11	11
Thermos carrier	11	11
Temperature chart for refrigerator	11	11
Refrigerator	9	11
Maternal health equipment		
Pinard obstetric stethoscope	11	11
Adult blood pressure cuff	11	11
Measuring tape	11	11
Adult scale	10	10
Vaginal speculum	10	11
Delivery table	9	9

Gynecological table	7	10
Pelvimeter	7	7
Gynecological lamp	2	3
Hemoglobin tester	0	5
Wheelchair[e]	–	6
Sterile equipment and materials		
Sterile gloves	11	11
Disinfectant	11	11
Suture	11	11
Disposable needles and syringes[d]	11	–
Disposable needles[e]	–	11
Disposable syringes[e]	–	11
Container for needles and medical waste	11	11
Wash drums	10	11
Wash bottles	7	7
Cold decontamination container	5	8
Trash can[e]	–	10
Sterile dressings[e]	–	11
Sink[e]	–	11
Soap[e]	–	11
Sterile towels/napkins[e]	–	7
Diagnostic and curative care equipment for children		
Stethoscope	11	11
Thermometer	11	11
Otoscope	10	11
Examination table	10	11
Child blood pressure cuff	3	1
Laryngeal mirror	1	3
Flashlight[e]	–	9
Suction bulb or electric suction device[e]	–	9
Metal tongue depressor[e]	–	8

[a] Values are counts (i.e., number of health centers). No statistical tests were conducted.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up ranged from N = 10 to 11.
[d] Information not collected at follow-up.
[e] Information not collected at baseline.

All health centers generally stocked key vaccines at both baseline and follow-up, but stock ruptures were common[36,37] (**Table 25**). At baseline, there had only been a rupture in the polio supply chain (at two health centers); however, at follow-up, at least one health center reported ruptures in the supply chain of every key vaccine except Bacille Calmette-Guérin (BCG) and measles.

Micronutrient and supplement availability was irregular in both 2010 and 2014. All health centers stocked 100,000 IU doses of vitamin A and ORS at both time points; however, the number of health centers that had 200,000 IU doses declined from 11 to 7 between 2010 and 2014. The number of health centers that stocked folic acid, iron syrup, and Plumpy'Nut also declined. At follow-up, there was a marginal improvement in the availability of iron folate (from 10 to 11 centers) and iron sulfate pills (from 8 to 9), and a more marked improvement in the availability of zinc (from 6 to 10). Multivitamins (in drinkable and pill form) and vitamin B complex were not commonly stocked at follow-up. In 2014, stock ruptures occurred more frequently for 200,000 IU doses of vitamin A, iron folate pills, and ORS, and nearly all health centers that stocked iron syrup reported shortages. The number of stock ruptures did not change

[36] A stock rupture was defined as not having the product available sometime during the preceding 6 months.
[37] Note that *Tubaramure*, under FFP guidelines, was not able to purchase immunizations or medications.

between surveys for folic acid pills. For all other micronutrients and supplements, the number of stock ruptures fell, most noticeably for 100,000 IU vitamin A, Plumpy'Nut, and zinc.

For the treatment of infections, antibiotics were generally available, and stock ruptures were infrequent. Amoxicillin, chloramphenicol, and either co-trimoxazole or erythromycin were universally available in 2010, and all were universally available in 2014. The availability of benzylpenicillin and gentamycin fell slightly between surveys. At follow-up, all but one health center stocked tetracycline ophthalmic ointment (used for eye infections and to prevent blindness from measles). The frequency of stock ruptures fell for chloramphenicol between 2010 and 2014, and no stock ruptures were reported for amoxicillin, chloramphenicol, and co-trimoxazole in 2014. No more than three health centers reported stock ruptures for erythromycin, benzylpenicillin and gentamycin (often used in combination), and tetracycline ophthalmic ointment at follow-up.

For the treatment of malaria, quinine as well as amodiaquine and artesunate were stocked in all health centers in 2010 and 2014. Stock ruptures fell between surveys for amodiaquine and artesunate so that no health center experienced a rupture in 2014. Although no health center experienced a stock rupture for quinine in 2010, one health center experienced a shortage in 2014. For deworming, either albendazole or mebendazole was universally available at all health centers at baseline and follow-up, and the number of centers reporting stock ruptures fell from one to zero. For the treatment of *schistosoma* and *taenia*, only one health center stocked praziquantel and niclosamide at follow-up (no information was collected at baseline), and stock ruptures occurred for both. Five of the 11 health centers were designated to stock antiretroviral medications for the prevention of mother-to-child HIV transmission in 2014.[38] Of these health centers, all five reported stocking lamivudine (with only one reporting a stock rupture), four stocked zidovudine (with no stock ruptures), and three stocked both tenofovir and nevirapine (with no stock ruptures).

With regard to analgesics, acetaminophen was universally available at both baseline and follow-up, and the availability of diazepam improved to universal availability at follow-up. Stock ruptures were infrequent at both surveys. There were no stock ruptures for acetaminophen in 2014, while the number of stock ruptures for diazepam increased from zero to two stock ruptures at follow-up. Also at follow-up, ibuprofen was universally available (no stock ruptures), diclofenac sodium was available at 10 health centers (four stock ruptures), and aspirin was available at 9 health centers (four stock ruptures). Oral contraceptives were stocked at all 11 health centers at follow-up, and no center experienced stock ruptures.

Table 25. Availability of medications and supplies[a]

	Baseline[b]		Follow-up[c]	
	Generally stocked	Stock rupture past 6 months	Generally stocked	Stock rupture past 6 months
N	**11**	**11**	**11**	**11**
Immunizations and distilled water				
Polio	11	2	11	2
DPT-HiB-HepB	11	0	11	1
BCG	11	0	11	0
Measles	11	0	11	0
Tetanus	11	0	11	1
Distilled water	11	0	11	1

[38] Ministry of Health guidelines include these antiretroviral drugs on their list of essential medications for health centers to provide. Their availability was being phased in, and at the time of the 2014 survey, only 5 of the 11 centers were designated as locations where these medications would be available.

Micronutrients and supplements				
100,000 IU vitamin A	11	4	11	0
200,000 IU vitamin A	11	2	7	3
Folic acid pills	10	2	8	2
Iron folate pills	10	1	11	2
Iron sulfate pills	8	2	9	1
Iron syrup	8	8	7	6
ORS	11	1	11	2
Plumpy'Nut	10	6	8	1
Zinc	6	4	10	1
Multivitamin pills[d]	–	–	5	2
Drinkable multivitamins[d]	–	–	1	1
Vitamin B complex[d]	–	–	3	1
Antibiotics				
Amoxicillin	11	0	11	0
Chloramphenicol	11	3	11	0
Co-trimoxazole or erythromycin[e]	11	0	–	–
Co-trimoxazole[d]	–	–	11	0
Erythromycin[d]	–	–	11	2
Benzylpenicillin and gentamycin[e]	9	3	–	–
Benzylpenicillin[d]	–	–	8	1
Gentamycin[d]	–	–	8	3
Tetracycline ointment[d]	–	–	10	2
Antifungal				
Nystatin[d]	–	–	11	1
Antimalarial				
Amodiaquine and artesunate	11	3	11	0
Quinine	11	0	11	1
Anthelminthic				
Albendazole or mebendazole	11	1	11	0
Praziquantel[d]	–	–	1	1
Niclosamide[d]	–	–	1	1
Antiretroviral[f]				
Lamivudine (3TC)[d]	–	–	5	1
Zidovudine (ZDV or AZT)[d]	–	–	4	0
Tenofovir (TDF)[d]	–	–	3	0
Nevirapine (NVP)[d]	–	–	3	0
Analgesic				
Acetaminophen	11	1	11	0
Diazepam	10	0	11	2
Ibuprofen[d]	–	–	11	0
Diclofenac sodium[d]	–	–	10	4
Aspirin[d]	–	–	9	4
Oral contraceptives				
Oral contraceptives generally stocked[d]	–	–	11	0

[a] Values are counts (i.e., number of health centers). No statistical tests were conducted.
[b] Sample size at baseline did not vary for "generally stocked." Sample size for "stock rupture" was the number of health centers that generally stocked the particular item.
[c] Sample size at follow-up did not vary for "generally stocked." Sample size for "stock rupture" was the number of health centers that generally stocked the particular item.
[d] Information not collected at baseline.
[e] Information not collected at follow-up.
[f] Only 5 of the 11 health centers were designated to provide antiretroviral medications in 2014.

4. Results: Household Characteristics

4.1 Household Demography and Housing

Household demographics were not meaningfully different between the 2010 and 2014 study samples. At both time points, households had on average of 6.0 members with about 3.8 members younger than 18 years of age (**Table 26**). At both surveys, nearly all of the household heads were male and had farming as their primary occupation. The majority of household heads had no education or incomplete primary schooling and were on average between 36 and 37 years old. Although education levels of household heads were low in both 2010 and 2014, they were somewhat higher in 2014, when the percentage without schooling dropped from 44.0 percent to 37.2 percent.

Housing conditions were similar at baseline and follow-up. Nearly all households lived in a house they owned and did not share the dwelling with another household (**Table 27**). In general, housing conditions were poor: houses had on average of three rooms, almost all had dirt floors, and most had walls made of adobe bricks (76.1 percent at baseline and 77.4 percent at follow-up) or rammed earth[39] (17.5 percent at baseline and 12.3 percent at follow-up). The only noticeable difference in housing conditions between 2010 and 2014 was that fewer households lived under a roof made out of thatch/straw in the 2014 study sample (38.5 percent) than in the 2010 study sample (47.1 percent).

A greater percentage of households had access to tap water in 2014 than in 2010 (81.5 percent and 74.2 percent, respectively), but the time to reach the water source was about the same—around 40 minutes—at both surveys (**Table 28**). Only a very small percentage of households had electricity at either survey point (0.4 percent at baseline and 1.1 percent at follow-up). However, almost all households used firewood or straw for cooking at both time points (97.8 percent at baseline and 95.4 percent at follow-up). Household light sources were considerably different between the baseline and follow-up study samples. Reported use of battery-operated equipment was nearly two-fold greater at follow-up (80.7 percent) than at baseline (44.3 percent), whereas the reported use of natural combustible material or oil products was more than two-thirds lower at follow-up (9.8 percent and 5.8 percent, respectively, at follow-up versus 30.6 percent and 20.8 percent, respectively, at baseline). The only (meaningful) differences between study arms at baseline were the average number of rooms in houses, the type of wall material, and the time to get drinking water.

4.2 Household Assets

Nearly all households owned a house and land at both baseline and follow-up (**Table 29**). At follow-up, households owned an average of 33 household goods (around seven more than at baseline) and five pieces of agricultural equipment (a 0.7-unit increase), two small animals (a 0.1-unit increase), and two medium-sized animals (a 0.1-unit increase); very few households owned large animals, a motorbike, or a bicycle at either baseline or follow-up. At both baseline and follow-up, the number of household goods, pieces of agricultural equipment, and animal ownership were different among study arms: the number of household goods and pieces of agricultural equipment were highest in the T24 group and lowest in the control group at both baseline and follow-up, ownership of large animals was lowest in the control group and highest in the T18 and TNFP groups at baseline, and ownership of medium animals was highest in the T24 group and lowest in the control group at follow-up.

[39] Rammed earth (*pisé* in French) is a technique for building walls using raw materials, such as earth, chalk, lime, and gravel. The damp material is poured and then compacted to construct the wall.

Table 26. Characteristics of households[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3563**	**2568**	**985**	**585**	**998**	**995**	**3560**	**2982**	**989**	**997**	**996**	**578**
Household												
Size	5.9 ± 2.0	5.9 ± 2.0	5.9 ± 2.0	5.9 ± 2.0	5.9 ± 1.9	5.9 ± 2.0	6.0 ± 2.0	6.0 ± 2.0	6.0 ± 1.9	6.0 ± 2.0	6.0 ± 1.9	5.9 ± 2.0
Members under 18 years	3.7 ± 1.7	3.7 ± 1.7	3.7 ± 1.7	3.7 ± 1.7	3.7 ± 1.7	3.7 ± 1.8	3.8 ± 1.7	3.8 ± 1.7	3.7 ± 1.7	3.8 ± 1.8	3.8 ± 1.7	3.8 ± 1.8
Members 18 years and older	2.2 ± 0.6	2.2 ± 0.6	2.2 ± 0.7	2.2 ± 0.6	2.2 ± 0.6	2.1 ± 0.6	2.2 ± 0.7	2.2 ± 0.7	2.2 ± 0.7	2.2 ± 0.7	2.2 ± 0.7	2.2 ± 0.6
Members under 60 months	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.5 ± 0.6	1.6 ± 0.6	1.6 ± 0.6	1.6 ± 0.6
Adults (%)	39.9 ± 12.7	40.0 ± 12.6	40.1 ± 12.9	40.1 ± 12.2	39.7 ± 12.5	39.6 ± 12.8	39.4 ± 12.7	39.4 ± 12.6	39.8 ± 12.7	39.2 ± 12.7	39.2 ± 12.4	39.5 ± 13.0
Dependency ratio	1.6 ± 0.8	1.5 ± 0.8	1.5 ± 0.8	1.5 ± 0.8	1.6 ± 0.8	1.6 ± 0.8	1.6 ± 0.8	1.6 ± 0.8	1.5 ± 0.8	1.6 ± 0.8	1.6 ± 0.8	1.6 ± 0.8
Household head												
Age	36.6 ± 10.5	36.7 ± 10.6	36.7 ± 10.7	37.0 ± 10.3	36.5 ± 10.6	36.3 ± 10.3	36.4 ± 10.5	36.5 ± 10.5	37.4 ± 11.0	36.1 ± 10.0	36.0 ± 10.5	35.9 ± 10.8
Male (%)	92.1	92.0	92.4	90.4	92.7	92.1	92.2	92.2	91.3	92.7	92.7	92.4
Household head education												
None/preschool	44.0	42.9	39.9	48.5	42.6	46.9	37.2	36.8	35.9	35.4	39.1	39.0
Primary incomplete	50.1	50.6	54.1	44.8	50.5	48.7	54.2	54.1	55.7	55.5	51.1	54.8
Primary complete	1.8	1.9	1.7	1.9	2.1	1.6	3.3	3.4	4.0	2.7	3.5	2.4
Secondary incomplete	3.9	4.3	4.1	4.5	4.5	2.7	5.1	5.3	4.2	5.7	5.9	3.8
Higher education	0.2	0.2	0.2	0.3	0.2	0.0	0.3	0.4	0.1	0.6	0.4	0.0
Household head occupation												
Unemployed	1.0	1.1	1.3	0.7	1.0	0.7	1.1	1.1	1.2	1.0	1.0	1.2
Farms own or family land	78.0	76.5	78.8	75.7	74.7	81.9	74.6	74.8	75.7	72.8	75.8	73.8
Farms someone else's land	2.8	2.5	1.9	2.9	2.9	3.6	5.8	5.0	6.3	5.6	3.0	10.2
Agricultural laborer	5.2	5.0	4.4	5.0	5.7	5.5	3.2	3.0	2.6	3.2	3.1	4.3
Retailer (e.g., has a store)[d]	0.4	0.4	0.4	0.2	0.4	0.4	0.4	0.4	0.1	0.5	0.5	0.3
Market/trade	2.4	2.5	2.4	3.4	2.0	2.0	3.0	3.0	2.0	4.4	2.5	2.8
Office/institution	4.5	5.5	4.6	6.0	6.0	1.8	4.1	4.5	3.3	4.1	6.0	2.1
Manual labor	4.7	5.2	5.2	4.3	5.7	3.2	5.8	6.1	6.5	6.3	5.6	4.0
Other	1.2	1.3	1.0	1.7	1.4	0.8	2.1	2.2	2.2	2.0	2.4	1.2

[a] Values are % or mean ± SD. SDs adjusted for clustering.

[b] Sample size at baseline ranged from N = 3547 to 3563 in the full sample; N = 2556 to 2568 in the treatment arms; N = 984 to 985 in the T24 arm; N = 580 to 585 in the T18 arm; N = 991 to 998 in the TNFP arm; and N = 990 to 995 in the control arm.

[c] Sample size at follow-up ranged from N = 3555 to 3560 in the full sample; N = 2978 to 2982 in the treatment arms; N = 988 to 989 in the T24 arm; N = 994 to 997 in the T18 arm; N = 995 to 996 in the TNFP arm; and N = 577 to 578 in the control arm.

[d] Retail is a more formal form of trade, involving keeping a premise or shop that is owned or rented. Market/trade is informal or petty trade, such as a market stall or street vending.

* There were no statistical differences among study arms, $p < 0.05$.

Table 27. Housing characteristics[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3560**	**2565**	**985**	**585**	**996**	**995**	**3560**	**2982**	**989**	**997**	**996**	**578**
Owns home	98.2	98.2	99.0	96.8	98.3	98.2	97.5	97.7	97.5	98.1	97.4	96.7
Dwelling type and characteristics												
Number of rooms	3.3 ± 1.2*	3.3 ± 1.2	3.5 ± 1.2	3.3 ± 1.2	3.2 ± 1.2	3.1 ± 1.2	3.5 ± 1.1	3.6 ± 1.1	3.6 ± 1.1	3.5 ± 1.2	3.6 ± 1.1	3.4 ± 1.1
Shared dwelling	0.5	0.5	0.4	0.5	0.5	0.6	0.8*	0.6	0.9	0.4	0.5	1.9
Housing quality												
Type of floor												
Dirt	96.2	95.4	97.2	94.9	94.0	98.2	94.0	93.4	94.4	94.9	90.8	97.6
Type of walls												
Rammed earth	17.5	13.2	7.1	17.4	16.8	28.5	12.3	10.1	4.7	13.4	12.1	24.1
Adobe bricks	76.1	79.8	89.2	76.8	72.4	66.6	77.4	79.3	89.5	77.1	71.3	67.7
Clay bricks	4.1	4.5	1.8	3.2	7.9	3.0	8.7	9.1	4.2	8.1	14.9	6.9
Cement bricks/stone	0.8	0.9	1.3	0.5	0.8	0.6	1.0	1.0	1.2	0.7	1.2	0.7
Other	1.4*	1.5	0.5	2.1	2.1	1.3	0.5*	0.5	0.4	0.6	0.6	0.5
Type of roof												
Thatch/straw	47.1	45.3	45.6	45.3	45.0	51.9	38.5	36.9	35.0	38.4	37.4	46.3
Corrugated aluminum	39.4	40.0	40.0	42.1	38.9	37.6	44.6	45.8	47.8	46.5	43.2	38.3
Concrete/tile	13.5	14.7	14.4	12.6	16.1	10.6	16.9	17.2	17.2	15.0	19.4	15.4

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3539 to 3560 in the full sample; N = 2549 to 2565 in the treatment arms; N = 978 to 985 in the T24 arm; N = 582 to 585 in the T18 arm; N = 989 to 996 in the TNFP arm; and N = 990 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3511 to 3560 in the full sample; N = 2941 to 2982 in the treatment arms; N = 980 to 989 in the T24 arm; N = 973 to 997 in the T18 arm; N = 988 to 996 in the TNFP arm; and N = 570 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

Table 28. Water and energy access[a]

	Baseline[b]						Follow-up[c]					
		Study arm					Full sample	Study arm				
	Full sample	All treated	T24	T18	TNFP	Control		All treated	T24	T18	TNFP	Control
N	3563	2568	985	585	998	995	3559	2981	989	997	996	578
Drinking water source												
Tap water	74.2	76.5	81.7	66.2	77.4	68.1	81.5	84.0	89.2	78.2	84.6	68.9
Open well	2.4	2.2	0.3	3.6	3.2	3.1	2.9	2.3	1.8	1.7	3.3	5.9
Covered well	1.7	1.4	1.3	0.5	1.9	2.7	1.3	1.3	0.9	1.1	1.9	1.4
Surface water	21.6	19.9	16.5	29.7	17.5	26.0	14.3	12.4	8.1	19.0	10.1	23.9
Time to get drinking water (minutes)	37.9 ± 37.4*	36.4 ± 33.6	30.3 ± 24.0	41.7 ± 42.3	39.3 ± 35.2	41.7 ± 45.6	40.9 ± 53.7	40.4 ± 54.6	31.3 ± 34.8	51.3 ± 78.7	38.6 ± 36.5	43.3 ± 48.6
Household has electricity	0.4	0.6	0.3	1.2	0.5	0.0	1.1	1.2	0.4	2.1	1.2	0.5
Energy for cooking												
Charcoal	1.9	2.2	1.1	2.1	3.2	1.1	4.1	4.8	3.5	4.4	6.4	0.7
Firewood/straw	97.8	97.5	98.7	97.8	96.2	98.7	95.4	94.7	95.9	95.0	93.1	98.9
Other	0.3	0.4	0.2	0.2	0.6	0.2	0.5	0.5	0.5	0.6	0.5	0.4
Energy for light												
Electricity	0.6	0.7	0.2	1.4	0.8	0.2	1.5	1.7	1.1	2.0	1.9	0.9
Kerosene/oil	20.8	24.7	30.5	19.8	21.9	10.7	5.8	6.4	9.8	4.6	4.7	2.8
Candle	3.8	4.5	4.5	4.3	4.7	2.0	2.2	2.5	2.6	1.7	3.1	0.5
Battery-operated equipment	44.3	40.9	40.2	43.2	40.1	53.0	80.7	80.3	77.9	82.0	81.1	82.7
Firewood/straw/coal/dung	30.6	29.2	24.6	31.3	32.5	34.1	9.8	9.1	8.6	9.5	9.1	13.1

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3538 to 3563 in the full sample; N = 2544 to 2568 in the treatment arms; N = 975 to 985 in the T24 arm; N = 580 to 585 in the T18 arm; N = 989 to 998 in the TNFP arm; and N = 993 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3482 to 3559 in the full sample; N = 2919 to 2981 in the treatment arms; N = 960 to 989 in the T24 arm; N = 972 to 997 in the T18 arm; N = 975 to 996 in the TNFP arm; and N = 563 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

Table 29. Asset ownership[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	3563	2568	985	585	998	995	3560	2982	989	997	996	578
Ownership (%)												
House(s)/apartment(s)	97.4	97.6	98.1	97.3	97.4	97.0	97.2	97.5	97.8	97.6	97.1	95.7
Plot(s) of land	97.6	98.1	98.3	97.6	98.1	96.6	97.4	97.8	97.8	98.6	97.0	95.7
Ownership (number)												
Household goods	25.4 ± 13.7*	26.6 ± 14.3	28.1 ± 13.2	25.2 ± 14.1	25.9 ± 15.3	22.2 ± 11.3	32.6 ± 15.9*	33.7 ± 16.1	34.7 ± 15.7	32.9 ± 15.2	33.6 ± 17.4	26.9 ± 13.3
Agricultural equipment	4.6 ± 2.4*	4.7 ± 2.6	5.0 ± 2.8	4.5 ± 2.4	4.6 ± 2.4	4.2 ± 2.1	5.3 ± 2.6*	5.4 ± 2.6	5.7 ± 2.7	5.3 ± 2.6	5.2 ± 2.6	4.8 ± 2.2
Small animals (chicken, rabbit)	2.1 ± 3.6	2.1 ± 3.8	2.0 ± 3.5	2.4 ± 4.0	2.0 ± 3.8	1.9 ± 3.3	2.2 ± 4.0	2.2 ± 3.9	2.2 ± 3.8	2.6 ± 4.5	1.8 ± 3.2	2.4 ± 4.2
Medium animals (goat, sheep)	1.6 ± 2.2	1.7 ± 2.2	1.7 ± 2.0	1.6 ± 2.4	1.7 ± 2.3	1.4 ± 2.1	1.7 ± 2.3*	1.8 ± 2.3	2.0 ± 2.4	1.7 ± 2.3	1.7 ± 2.2	1.4 ± 2.1
Large animals (cow, pig)	0.3 ± 1.0*	0.4 ± 1.1	0.3 ± 0.8	0.4 ± 1.2	0.4 ± 1.3	0.2 ± 0.7	0.5 ± 1.8	0.5 ± 1.9	0.4 ± 1.1	0.6 ± 2.8	0.5 ± 1.2	0.5 ± 1.1
Motorbike/bike	0.4 ± 0.6	0.4 ± 0.7	0.4 ± 0.8	0.4 ± 0.5	0.4 ± 0.5	0.4 ± 0.5	0.4 ± 0.5	0.4 ± 0.5	0.4 ± 0.5	0.5 ± 0.5	0.4 ± 0.6	0.4 ± 0.5

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
* Study arms differed, $p < 0.05$.

5. Results: Maternal Characteristics

Mothers were on average 31 years old at baseline and 30 years old at follow-up, and nearly all were the biological mother of the index child (**Table 30**). The percentage of mothers who had a spouse or partner was lower at follow-up than at baseline (81 percent and 89 percent, respectively). Overall, maternal education was very low: only half had ever attended school, and of those that had, very few had completed primary school. The percentage of mothers who had never attended school was lower in the follow-up sample than in the baseline sample (46 percent and 54 percent, respectively). Literacy rates increased, but only slightly, from 52.2 percent at baseline to 54.2 percent at follow-up. At baseline, the age of mothers varied across the study arms.

At both baseline and follow-up, more than 90 percent of mothers reported having worked in the past 12 months. Similar to household heads, almost all reported working in farming and agriculture. The percentage not remunerated for their work increased from 36.1 percent at baseline to 44 percent at follow-up. The percentage of mothers who perceived contributing all or almost all of the resources for household expenses was higher at follow-up than at baseline (23.8 percent and 13.3 percent, respectively). Whether a mother worked during the previous 12 months was significantly different across study arms at baseline.

Table 30. Maternal characteristics and activities[a]

	Baseline[b]						Follow-up[c]					
	Full sample	Study arm					Full sample	Study arm				
		All treated	T24	T18	TNFP	Control		All treated	T24	T18	TNFP	Control
N	3563	2568	985	585	998	995	3560	2982	989	997	996	578
Age	30.9 ± 7.8*	31.0 ± 7.8	31.3 ± 7.9	31.4 ± 8.3	30.6 ± 7.3	30.5 ± 7.9	30.2 ± 7.2*	30.3 ± 7.2	31.0 ± 7.7	30.0 ± 7.0	30.0 ± 6.9	29.4 ± 6.8
Married	88.6	89.2	90.5	87.3	89.0	87.0	80.8	81.9	84.8	78.9	82.0	74.9
Biological mother	97.4	97.5	97.2	96.9	98.1	97.1	98.2*	98.0	97.1	98.1	98.8	99.1
Education												
None/preschool	53.9	53.2	48.7	55.7	56.1	55.7	46.1	45.0	44.7	42.3	48.0	51.5
Primary incomplete	42.1	42.0	47.2	38.4	39.1	42.3	48.2	48.8	48.5	51.6	46.2	45.6
Primary complete	1.2	1.4	1.3	1.7	1.4	0.6	1.9	2.0	2.7	1.8	1.6	1.0
Secondary incomplete	2.7	3.2	2.7	3.9	3.3	1.4	3.7	4.1	4.0	4.2	4.0	1.9
Higher education	0.1	0.1	0.0	0.2	0.1	0.0	0.1	0.1	0.0	0.1	0.1	0.0
Literacy												
Literate	52.2	53.2	56.2	49.9	52.1	49.5	54.2	55.7	58.0	56.3	52.8	46.4
Partially literate	9.5	10.4	11.6	10.9	8.9	7.2	9.3	8.9	9.3	7.6	9.7	11.2
Illiterate	38.3	36.4	32.2	39.1	39.0	43.2	36.5	35.4	32.7	36.0	37.4	42.4
Work during the past 12 months												
None	9.0	7.8	6.8	7.7	8.7	12.3	7.1	7.2	6.5	5.0	10.0	6.4
Yearlong	69.8	69.4	72.2	68.3	67.4	70.6	71.9	71.7	72.4	73.7	68.9	73.0
Seasonal	8.8	10.8	9.8	12.6	10.7	3.8	7.7	8.0	6.1	9.8	8.0	6.3
Sometimes	12.4*	12.0	11.2	11.4	13.2	13.3	13.4	13.2	15.1	11.5	13.1	14.3
Main occupation												
Unemployed	9.0	7.7	6.7	7.7	8.6	12.2	6.9	7.0	6.4	4.9	9.8	6.2
Farms own or family land	79.7	81.4	84.1	80.3	79.4	75.4	79.6	80.3	81.3	80.9	78.6	76.3
Farms someone else's land	3.7	3.2	3.3	3.1	3.1	5.0	6.7	5.9	6.2	7.7	3.7	11.1
Agricultural labor	5.0	4.5	3.2	5.6	5.2	6.4	2.4	2.1	1.9	2.0	2.4	4.0
Retail	0.1	0.1	0.1	0.0	0.1	0.0	0.2	0.2	0.2	0.1	0.2	0.2
Market/trade	0.6	0.6	0.7	0.7	0.4	0.6	1.8	1.9	2.2	1.7	1.9	1.0
Office/institution	1.4	1.8	1.5	1.9	2.0	0.3	1.7	2.0	1.2	2.2	2.5	0.5
Manual labor	0.4	0.6	0.2	0.3	1.1	0.0	0.3	0.3	0.2	0.1	0.6	0.3
Other	0.1	0.2	0.2	0.3	0.0	0.1	0.3	0.3	0.3	0.3	0.2	0.3
Earnings												
Cash	7.1	7.3	5.0	8.0	9.3	6.6	10.9	11.0	9.9	9.3	14.1	10.5
In-kind	30.9	32.4	33.9	31.9	31.1	27.0	27.2	26.5	24.6	25.5	29.5	30.8
Cash and in-kind	25.8	27.0	28.4	25.0	26.9	22.3	17.5	17.1	17.3	17.1	16.7	19.9
Other compensation	0.1	0.1	0.1	0.0	0.2	0.0	0.4	0.4	0.7	0.4	0.0	0.4
Nothing	36.1	33.1	32.5	35.2	32.5	44.1	44.0	45.1	47.6	47.7	39.7	38.4
Mother's perceived contribution to household expenses												
Nothing	23.4	23.1	25.1	23.6	20.8	24.3	15.4	16.0	17.0	15.7	15.2	12.7
Almost nothing	15.1	15.4	17.9	13.4	14.0	14.6	12.9	12.9	11.0	11.5	16.2	12.5
A little	48.1	48.4	44.4	46.7	53.4	47.5	47.9	48.4	47.2	51.7	46.2	45.8
All/almost all	13.3	13.2	12.7	16.3	11.8	13.6	23.8	22.7	24.8	21.1	22.4	29.1

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3245 to 3563 in the full sample; N = 2371 to 2568 in the treatment arms; N = 919 to 985 in the T24 arm; N = 540 to 585 in the T18 arm; N = 912 to 998 in the TNFP arm; and N = 874 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3307 to 3560 in the full sample; N = 2765 to 2982 in the treatment arms; N = 923 to 989 in the T24 arm; N = 946 to 997 in the T18 arm; N = 896 to 996 in the TNFP arm; and N = 542 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

6. Results: Child Characteristics

Index children were on average between 32 and 33 months of age at baseline and 33 and 34 months of age at follow-up (**Table 31**). At baseline and follow-up, 48.3 percent and 51.1 percent of children, respectively, were boys. There were no differences at baseline across study arms in the average age of the index children or the percentage who were male.

Table 31. Child characteristics[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3563**	**2568**	**985**	**585**	**998**	**995**	**3560**	**2982**	**989**	**997**	**996**	**578**
Age (months)	32.8 ± 5.6	32.8 ± 5.6	32.7 ± 5.	32.7 ± 5.	33.0 ± 5.6	32.8 ± 5.6	33.2 ± 5.3	33.2 ± 5.3	33.1 ± 5.4	33.3 ± 5.4	33.2 ± 5.	33.4 ± 5.3
Sex (boys)	48.3	47.7	46.6	46.2	49.8	49.7	51.1	51.1	49.5	52.9	50.9	51.0

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up did not vary.
* There were no statistical differences among study arms, $p < 0.05$.

7. Results: *Tubaramure* Past Participation

7.1 Awareness and Past Participation in *Tubaramure*

At follow-up, approximately 92 percent of respondents in treatment *collines* were aware of *Tubaramure,* and 72 percent claimed that they had previously participated[40] (**Table 32**). Both the program awareness and the beneficiary status differed significantly among respondents in the control and treatment groups: the control group was less likely to be aware of the program and to have been a beneficiary, as would be expected. Approximately two-fifths of the control group respondents were unaware of *Tubaramure*; however, a small percentage (4.2 percent) claimed that they were past program beneficiaries. The T18 group had fewer past beneficiaries than other treatment groups, by 5 to 10 percentage points. Among those who reported that they were past beneficiaries and lived in treatment *collines*, 17.6 percent were *Tubaramure* leader mothers. Almost all past beneficiaries were aware that *Tubaramure* provided food rations (97 percent), and 83 percent also mentioned BCC as a program activity. There was more variation between treatment arms in the percentage of respondents who were aware that *Tubaramure* provided cooking demonstrations. Across study arms, less than half of the past beneficiaries mentioned cooking demonstrations as a program activity, and this varied from 42 percent in the T24 group to 51 percent in the TNFP group.

7.2 Past Participation in *Tubaramure* Program Activities

Almost all of the past beneficiary mothers (of children 24–41 months) interviewed reported receiving food rations during their participation in the *Tubaramure* program (93 percent to 96 percent) (**Table 33**). However, only about three-fourths of those in the T24 and TNFP groups and 69 percent in the T18 group reported participating in care groups. Care groups were expected to meet every 2 weeks or twice a month. Approximately 45 percent of beneficiary mothers in the T24 and TNFP groups reported attending a care group more than once a month. Although slightly fewer beneficiary mothers in the T18 group than in other treatment arms reported ever participating in care groups, a higher percentage of those in the T18 group reported attending care groups more than once a month (55.4 percent). The results indicate that attendance by more than half of program participants in the treatment arms was lower than intended by the program design. Among those who had participated in a care group, 75.8 percent reported that a cooking demonstration had ever been offered. A little more than half of beneficiary respondents had received at least one visit from a *Tubaramure* leader mother. Those in the control arm did not participate in any activities because the program was not available to them.

[40] This is around 8 percentage points lower than what was reported in the 2012 survey (when the program was ongoing)(Leroy et al. 2014). The lower reported (past) participation in the 2014 survey might have been due to migration (which would have reduced the percentage of respondents in the treatment *collines* who were exposed to the program) and recall bias (i.e., former beneficiaries had forgotten the name of the program).

Table 32. *Tubaramure* program awareness and past participation[a]

	Follow-up[b]					
		Study arm				
	Full sample	All treated	T24	T18	TNFP	Control
N[c]	**3560**	**2982**	**989**	**997**	**996**	**578**
Aware of *Tubaramure*	86.8*	91.6	91.0	90.4	93.5	62.0
Beneficiary status						
Past beneficiary	61.0*	72.0	71.8	67.2	76.9	4.2
Never a beneficiary	39.0	28.0	28.2	32.8	23.1	95.8
Leader mother[c]	17.6	17.6	17.2	18.0	17.7	–
Still has a beneficiary card[c]						
Yes, presented	0.9	0.9	1.3	0.4	0.9	–
Yes, not presented	4.9	4.9	5.6	5.4	3.9	–
No	94.2	94.2	93.1	94.2	95.2	–
Mentioned awareness of program component[c]						
Rations	96.8	96.8	96.9	97.2	96.3	–
BCC	82.6	82.6	81.8	83.1	82.8	–
Cooking demonstrations	46.1	46.1	42.4	44.3	51.2	–

[a] Values are %.
[b] Sample size at follow-up ranged from N = 2090 to 3560 in the full sample; N = 2090 to 2982 in the treatment arms; N = 687 to 989 in the T24 arm; N = 657 to 997 in the T18 arm; N = 746 to 996 in the TNFP arm; and N = 0 to 578 in the control arm.
[c] Sample limited to past *Tubaramure* participants.
* Study arms differed, $p < 0.05$.

Table 33. *Tubaramure* program experiences (among past beneficiaries)[a]

	Follow-up[b]					
		Study arm				
	Full sample	All treated	T24	T18	TNFP	Control
N[c]	**2144**	**2144**	**709**	**669**	**766**	**–**
Ever received food ration	93.9	93.9	93.2	95.8	93.0	–
Ever participated in care group	74.3	74.3	76.3	68.9	77.0	–
Care group attendance						
More than once a month	47.6	47.6	43.6	55.4	45.2	–
Once a month	47.8	47.8	51.2	39.4	51.2	–
Less than once a month	4.6	4.6	5.2	5.3	3.6	–
Care group ever offered cooking demo	75.8	75.8	74.7	76.5	76.2	–
Leader mother ever visited home						
Yes	55.6	55.6	54.5	56.0	56.3	–
Interviewee herself is a leader mother	8.5	8.5	6.7	9.1	9.7	–

[a] Values are %.
[b] Sample size at follow-up ranged from N = 1581 to 2144 in the full sample; N = 1581 to 2144 in the treatment arms; N = 538 to 709 in the T24 arm; N = 457 to 669 in the T18 arm; and N = 583 to 766 in the TNFP arm.
[c] Sample limited to past *Tubaramure* participants.
* There were no statistical differences among study arms, $p < 0.05$.

8. Results: Household Post-Program Impact

8.1 Household Hygiene and Sanitation

The vast majority of households (97.3 percent at baseline and 94.1 percent at follow-up) did not treat their drinking water, even though nearly one-quarter of households used unsafe drinking water (see Section 4) (**Table 34)**. The percentage of households that covered their drinking water when storing it increased in all study arms, from 76.9 percent at baseline to 89.7 percent at follow-up. Garbage disposal practices changed considerably between the surveys: composting increased from about 50 percent to 68 percent, while the percentage discarding or burying garbage dropped substantially (from 46.0 percent to 33.2 percent and from 11.8 percent to 3.2 percent, respectively). At baseline, the percentage of households composting to dispose of garbage was significantly different between study arms.

Bed net ownership and use among all households increased dramatically between baseline and follow-up. The percentage of households that reported having a bed net increased from 54.7 percent at baseline to 97.4 percent at follow-up, and the mean number of bed nets owned by these households also increased from 2.1 at baseline to 2.7 at follow-up. Households owning bed nets were also more likely to report using them at follow-up than at baseline:[41] at baseline, 41.3 percent of households reported that at least one member slept under a bed net, and 34.5 percent reported that all members slept under bed nets; at follow-up, these percentages more than doubled, to 84.2 percent and 77.1 percent, respectively. Significant differences between study arms at baseline were found for all of the bed net outcomes.

Tubaramure had a post-program impact on the percentage of households treating drinking water (a 5.6 percentage point effect) in the T18 group (**Table 35**). The program had a marginally significant overall post-program effect on the percentage of households covering their drinking water (6.9 percentage points, p=0.053); it had a significant 8.4 percentage point effect in the T24 group. The program had a large (14 to 21 percentage point) post-program effect on the percentage of households that composted their trash (significant for all groups combined and in all three treatment groups separately). Finally, the results suggest that the program had a negative post-program effect on bed net ownership. At baseline, bed net ownership was considerably lower in the control group, but by the 2014 follow-up, ownership was nearly universal in all study arms (**Table 34**); over time, ownership in the control group thus rose faster than in the treatment groups. Note that at the time of the first follow-up in 2012, no program effect was found for bed net ownership (Leroy et al. 2014). The change thus appears to be due to factors unrelated to the *Tubaramure* program.

The majority of households (74.6 percent at baseline and 86.3 percent at follow-up) had soap available, and an increasing number of mothers reported using it on the day of the interview or the previous day (74.4 percent to 88.9 percent) (**Table 36**). Despite owning soap, using soap at key handwashing times was rare but generally increased over time. Soap ownership and use were different between study arms at baseline. The program had an overall positive post-program effect (5.5 percentage points) on the percentage of mothers washing hands with soap after defecating (significant in the T24 group). The only other post-program effect was found in the TNFP arm for washing hands after cleaning a child who had defecated (**Table 37**).

At both baseline and follow-up, only a minority of mothers (21.6 percent and 29.7 percent) and children (11.2 percent and 14.8 percent) were considered "clean" in a spot check of hands, hair, clothes, and face (**Table 38**). Among homesteads, around 32 percent of exteriors and 5 percent of interiors were considered

[41] Peak malaria season in Burundi is from November to March. Thus, reported bed net use in this survey (conducted from October to December) was most likely at its highest.

clean at both time points. There were no differences between study arms at baseline. Of the four domains of cleanliness, significant post-program improvements attributed to *Tubaramure* were seen only in the percentage of households with clean exteriors (improvement of 7.7 percentage points in all treatment groups combined and 9.3 percentage points in the TNFP group) (**Table 39**).

8.2 Household Food Security, Hunger, and Dietary Diversity

Food insecurity, as measured by the HFIAS, was prevalent at follow-up: more than 40 percent of households were severely food insecure and more than 37 percent moderately food insecure (**Table 40**). About 9 percent were experiencing severe hunger (measured by HHS, which uses the last three items of the HFIAS and reflects the most severe food insecurity experience) and 37 percent moderate hunger at baseline, but these percentages dropped to around 6 percent and 16 percent, respectively, at follow-up. Household dietary diversity improved from baseline to follow-up: households increased their consumption of different food groups during the previous day from 4.0 on average (out of a possible 12) at baseline to 4.8 at follow-up, and the percentage of households reporting having consumed food from fewer than four food groups dropped from around 65 percent to 44 percent.

Tubaramure had a post-program effect on food insecurity[42] (as measured by the HFIAS), reducing it by an average of 2 to 3 units (on a scale from 0 to 27; the effect was statistically significant in all three treatment groups combined and in each group separately (**Table 41**). The percentage of severely food insecure households decreased by 10 to 15 percentage points in each of the treatment groups (statistically significant in the TNFP group). Likewise, the percentage of food secure households increased by 4 to 7 percentage points in each treatment group (statistically significant in all three treatment groups). The post-program effect on hunger was limited to the TNFP arm (a 4.9 percentage point effect on the prevalence of severe hunger). *Tubaramure* also led to a moderate, statistically significant post-program increase in the number of food groups consumed (0.3 groups), and to an overall decline (around 12 percentage points) in the percentage of households with a dietary diversity score less than 4; the latter decline was also significant in the T18 (11.0 percentage points) and TNFP (19.6 percentage points) treatment arms.

[42] Note that these simple-difference impact estimates could be due to pre-existing differences between study arms. The double-difference impact estimates on other related outcomes in Table 41, however, indicate that the effect on food insecurity is plausible.

Table 34. Hygiene and sanitation[a]

	Baseline[1]						Follow-up[c]					
		Study arm						Study arm				
	Full sample[3]	All treated	T24	T18	TNFP	Control	Full sample[3]	All treated	T24	T18	TNFP	Control
N	3563	2568	985	585	998	995	3560	2982	989	997	996	578
Drinking water treatment method												
Nothing	97.3	97.3	97.1	98.1	97.2	97.1	94.1	93.7	95.7	91.6	93.7	96.4
Boiling	2.2	2.3	2.3	1.7	2.5	2.2	4.9	5.2	3.8	6.6	5.2	3.3
Other	0.5	0.4	0.6	0.2	0.3	0.7	1.0	1.1	0.4	1.8	1.1	0.3
Drinking water storage												
Uncovered container	22.7	23.1	22.9	22.3	23.7	21.7	9.8	8.9	7.4	8.3	11.0	14.5
Covered container	76.9	76.6	76.9	77.2	76.0	77.8	89.7	90.5	91.7	91.7	88.1	85.5
No storage	0.4	0.3	0.2	0.5	0.3	0.5	0.5	0.6	0.9	0.0	0.9	0.0
Garbage disposal[d]												
Discarded in a public space	46.0	44.5	39.5	47.7	47.5	49.8	33.2*	29.6	28.4	28.5	32.0	51.5
Burned	1.6	1.7	1.5	1.2	2.1	1.3	0.3	0.3	0.3	0.4	0.1	0.2
Buried	11.8	10.9	11.2	8.7	12.0	14.0	3.2	2.9	2.3	2.3	4.0	4.7
Composted	49.8*	52.6	57.1	49.6	49.8	42.7	67.9*	72.2	74.0	73.4	69.3	45.6
Fed to pigs/animals	0.5	0.5	0.2	0.9	0.6	0.5	0.3	0.2	0.1	0.5	0.1	0.3
Bed nets												
Households with bed nets	54.7*	58.0	60.0	56.1	57.2	46.2	97.4	97.7	97.8	97.0	98.2	96.0
If yes, number of bed nets	2.1 ± 1.0*	2.2 ± 1.0	2.2 ± 0.9	2.2 ± 1.0	2.2 ± 1.0	2.0 ± 0.9	2.7 ± 1.0	2.7 ± 1.0	2.7 ± 0.9	2.7 ± 0.9	2.6 ± 1.0	2.6 ± 1.0
Households that used a bed net previous night[e]	41.3*	44.2	48.9	42.3	40.8	33.6	84.2	85.2	85.3	83.9	86.5	78.9
All household members slept under a bed net previous night[e]	34.5*	37.3	41.4	35.5	34.4	27.2	77.1*	78.4	79.2	75.6	80.3	70.7

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 1933 to 3563 in the full sample; N = 1478 to 2568 in the treatment arms; N = 587 to 985 in the T24 arm; N = 326 to 585 in the T18 arm; N = 565 to 998 in the TNFP arm; and N = 455 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3454 to 3560 in the full sample; N = 2901 to 2982 in the treatment arms; N = 964 to 989 in the T24 arm; N = 964 to 997 in the T18 arm; N = 973 to 996 in the TNFP arm; and N = 553 to 578 in the control arm.
[d] Households reported all garbage disposal methods, thus totals add up to more than 100%.
[e] Calculated for all households, irrespective of bed net ownership.
* Study arms differed, $p < 0.05$.

Table 35. Hygiene and sanitation: post-program impact[a]

	Post-program impact[b]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Drinking water treatment method				
Nothing	-2.8 ± 1.8	-0.7 ± 2.2	-5.6 ± 2.1*	-2.9 ± 1.8
Drinking water storage				
Uncovered container	-6.9 ± 3.5	-8.4 ± 4.0*	-6.9 ± 3.8	-5.4 ± 4.0
Garbage disposal[c]				
Discarded in a public space	-16.5 ± 5.8*	-12.4 ± 6.4	-21.4 ± 6.7*	-17.0 ± 7.3*
Burned	-0.3 ± 0.6	-0.1 ± 0.8	0.3 ± 0.7	-0.9 ± 0.7
Buried	1.2 ± 3.3	0.1 ± 3.7	2.8 ± 3.8	1.1 ± 3.5
Composted	16.8 ± 5.1*	14.0 ± 5.9*	21.3 ± 5.6*	16.3 ± 6.1*
Fed to pigs/animals	-0.2 ± 0.4	0.0 ± 0.4	-0.2 ± 0.6	-0.4 ± 0.4
Bed nets				
Households with bed nets	-10.1 ± 2.8*	-12.0 ± 4.0*	-8.9 ± 4.3*	-9.0 ± 3.4*
If yes, number of bed nets	-0.2 ± 0.1	-0.2 ± 0.1	-0.1 ± 0.1	-0.2 ± 0.1
Households that used a bed net previous night[d]	-4.0 ± 3.9	-8.8 ± 5.1	-3.7 ± 4.2	0.4 ± 4.7
Household members who slept under a bed net previous night[d]	-2.1 ± 3.5	-5.6 ± 4.4	-3.5 ± 3.9	2.3 ± 4.3

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 5387 to 7123.
[c] Households reported all methods of garbage disposal used, thus totals add up to more than 100%.
[d] These percentages were calculated for all households, irrespective of having a bed net.
* Impact estimate significantly different from 0, $p < 0.05$.

Table 36. Soap use[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	3561	2567	985	585	997	994	3558	2982	989	997	996	576
Soap; mothers who:												
Have available in household	74.6*	76.1	80.4	70.3	75.3	70.5	86.3*	87.5	87.7	86.2	88.6	80.2
Used it today or yesterday	74.4*	76.3	81.9	71.8	73.3	69.7	88.9*	89.9	90.8	88.4	90.7	83.3
When used soap today or yesterday; mothers who washed:												
Child's hands	9.5	10.6	12.5	7.5	10.5	6.8	11.0	11.7	10.1	11.4	13.7	7.3
Own hands after defecation	4.6	4.9	3.8	5.1	5.8	3.8	13.8*	14.8	15.1	14.1	15.4	8.2
Own hands after cleaning a child that has defecated	4.3	4.7	5.2	5.1	4.0	3.2	9.1*	9.7	10.0	6.7	12.3	5.9
Own hands before feeding child	7.7	8.2	9.1	8.0	7.4	6.2	15.6	16.0	13.5	12.2	22.3	13.5
Own hands before preparing food	6.4	6.6	6.5	6.8	6.6	5.8	14.5	14.9	14.2	14.3	16.2	12.5
Own hands before eating	20.7	22.0	21.9	20.5	22.9	17.3	35.3	36.7	35.3	40.9	34.0	28.3

[a] Values are %.
[b] Sample size at baseline did not vary.
[c] Sample size at follow-up ranged from N = 3555 to 3558 in the full sample; N = 2979 to 2982 in the treatment arms; N = 988 to 989 in the T24 arm; N = 996 to 997 in the T18 arm; and N = 995 to 996 in the TNFP arm.
* Study arms differed, $p < 0.05$.

Table 37. Soap use: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Soap; mothers who:				
Have available in household	2.3 ± 2.7	-2.3 ± 3.3	6.5 ± 4.3	3.7 ± 3.0
Used it today or yesterday	0.4 ± 2.7	-4.7 ± 3.0	3.0 ± 3.8	3.7 ± 3.1
When used soap today or yesterday; mothers who washed:				
Child's hands	1.0 ± 3.0	-2.6 ± 3.8	3.4 ± 4.4	2.8 ± 3.3
Own hands after defecation	5.5 ± 2.4*	6.8 ± 3.1*	4.3 ± 3.4	5.1 ± 3.1
Own hands after cleaning a child that has defecated	2.6 ± 2.2	2.2 ± 2.5	-0.9 ± 2.7	5.6 ± 2.7*
Own hands before feeding child	1.3 ± 3.5	-2.4 ± 3.5	-3.0 ± 3.9	8.1 ± 4.9
Own hands before preparing food	1.9 ± 3.6	1.4 ± 4.8	0.7 ± 3.8	3.4 ± 5.2
Own hands before eating	3.3 ± 3.8	2.4 ± 3.9	8.7 ± 5.2	0.3 ± 5.6

[a] Values are double-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 7116 to 7119. One-sided tests are reported for all indicators.
[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.
* Impact estimate significantly different from 0, p < 0.05.

Table 38. Spot check observations[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3558**	**2564**	**984**	**585**	**996**	**994**	**3558**	**2980**	**989**	**996**	**996**	**578**
Mothers all clean	21.6	21.2	19.4	19.8	23.7	22.5	29.7	29.6	28.8	30.4	29.6	30.3
Children all clean	11.2	11.1	10.7	11.8	11.1	11.4	14.8	14.9	17.3	14.6	13.0	14.2
Exteriors all clean	31.9	32.3	31.1	35.4	31.6	30.9	32.4*	33.9	31.4	36.1	34.3	24.3
Interiors all clean[d]	4.3	4.7	5.6	6.0	3.2	3.2	5.5	5.2	3.1	3.4	9.0	7.5

[a] Values are %.
[b] Sample size at baseline ranged from N = 1636 to 3558 in the full sample; N = 1172 to 2564 in the treatment arms; N = 377 to 984 in the T24 arm; N = 300 to 585 in the T18 arm; N = 495 to 996 in the TNFP arm; and N = 464 to 994 in the control arm.
[c] Sample size at follow-up ranged from N = 1366 to 3558 in the full sample; N = 1192 to 2980 in the treatment arms; N = 353 to 989 in the T24 arm; N = 441 to 996 in the T18 arm; N = 398 to 996 in the TNFP arm; and N = 174 to 578 in the control arm.
[d] It was often not possible to observe the interior of the house, which led to a large number of missing values.
* Study arms differed, p < 0.05.

Table 39. Spot check observations: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Mothers all clean	0.5 ± 3.6	1.4 ± 4.8	2.5 ± 5.0	-1.9 ± 4.5
Children all clean	1.0 ± 2.6	3.8 ± 3.8	-0.2 ± 3.2	-0.9 ± 3.0
Exteriors all clean	7.7 ± 3.2*	6.9 ± 4.7	6.8 ± 5.2	9.3 ± 3.5*
Interiors all clean[d]	-2.5 ± 4.3	-6.5 ± 4.1	-5.4 ± 4.2	3.1 ± 5.3

[a] Values are double-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 3002 to 7116.
[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.
[d] It was often not possible to observe the interior of the house, which led to a large number of missing values.
* Impact estimate significantly different from 0, $p < 0.05$.

Table 40. Household hunger and dietary diversity[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	3563	2568	985	585	998	995	3560	2982	989	997	996	578
HFIAS												
Score (range 0 to 27)	–	–	–	–	–	–	11.1 ± 7.2*	10.7 ± 7.2	10.5 ± 7.4	11.1 ± 7.1	10.6 ± 7.0	13.2 ± 7.2
Category												
Food secure	–	–	–	–	–	–	11.8	12.6	13.6	11.4	12.8	7.5
Mildly insecure	–	–	–	–	–	–	7.7	8.3	10.2	7.9	6.7	4.7
Moderately insecure	–	–	–	–	–	–	37.1	37.8	35.8	36.8	40.7	33.7
Severely insecure	–	–	–	–	–	–	43.4	41.3	40.4	43.8	39.9	54.2
HHS												
Score (range 0 to 6)	1.4 ± 1.5	1.4 ± 1.5	1.3 ± 1.5	1.5 ± 1.5	1.5 ± 1.5	1.6 ± 1.5	0.8 ± 1.4	0.7 ± 1.3	0.7 ± 1.3	0.8 ± 1.3	0.6 ± 1.3	1.1 ± 1.6
Category												
Little-to-no hunger	54.0	56.1	59.8	54.4	53.5	48.5	78.2	79.7	80.1	78.2	80.7	70.8
Moderate hunger	37.3	35.6	32.6	37.4	37.5	41.5	15.7	15.1	14.4	15.9	15.1	18.9
Severe hunger	8.8	8.3	7.6	8.2	9.0	10.0	6.0	5.2	5.5	5.9	4.2	10.2
HDDS												
Score (range of 0 to 12)	4.0 ± 1.7	4.1 ± 1.7	4.1 ± 1.7	4.1 ± 1.6	4.1 ± 1.7	3.8 ± 1.6	4.8 ± 1.6*	4.9 ± 1.6	4.7 ± 1.6	4.9 ± 1.6	5.1 ± 1.6	4.3 ± 1.5
Low diversity (HDDS < 4)	64.9	63.7	62.6	62.4	65.4	68.2	44.0*	41.3	47.7	41.0	35.2	57.8

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3551 to 3563 in the full sample; N = 2559 to 2568 in the treatment arms; N = 982 to 985 in the T24 arm; N = 992 to 998 in the TNFP arm; and N = 992 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3542 to 3560 in the full sample; N = 2967 to 2982 in the treatment arms; N = 984 to 989 in the T24 arm; N = 989 to 997 in the T18 arm; N = 994 to 996 in the TNFP arm; and N = 575 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

Table 41. Household hunger and dietary diversity: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
HFIAS				
Score (range 0 to 27)	-2.5 ± 0.9*	-2.7 ± 1.0*	-2.1 ± 0.9*	-2.7 ± 0.9*
Category				
Food secure	5.1 ± 1.4*	6.1 ± 1.8*	4.0 ± 1.8*	5.3 ± 2.4*
Mildly insecure	3.6 ± 1.5*	5.5 ± 1.6*	3.2 ± 1.9	2.0 ± 1.9
Moderately insecure	4.1 ± 5.9	2.2 ± 6.5	3.2 ± 6.0	7.0 ± 7.3
Severely insecure	-12.8 ± 6.0*	-13.8 ± 6.9	-10.4 ± 6.2	-14.3 ± 6.5*
HHS				
Score (range 0 to 6)	-0.2 ± 0.1	0.0 ± 0.2	-0.2 ± 0.2	-0.3 ± 0.2*
Category				
Little-to-no hunger	1.4 ± 3.5	-2.0 ± 4.3	1.4 ± 4.5	4.6 ± 4.3
Moderate hunger	2.0 ± 3.1	4.4 ± 3.8	1.0 ± 4.2	0.3 ± 3.7
Severe hunger	-3.3 ± 2.2	-2.4 ± 2.6	-2.4 ± 2.5	-4.9 ± 2.4*
HDDS				
Score (range 0 to 12)	0.3 ± 0.2*	0.1 ± 0.2	0.3 ± 0.2	0.6 ± 0.2*
Low diversity (HDDS < 4)	-11.7 ± 5.8*	-4.2 ± 6.3	-11.0 ± 6.3*	-19.6 ± 7.3*

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points when data from both surveys were available and simple-difference impact estimates ± SE when only follow-up data were available. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 3542 to 7123.
[c] One-sided tests were conducted. The a priori hypothesis was a positive effect for HDDS score and negative effects for HFIAS score, HHS score, and low diversity (HDDS < 4).
* Impact estimate significantly different from 0, $p < 0.05$.

9. Results: Maternal Post-Program Impact

9.1 Feeding, Care, and Health Knowledge of Mothers

Overall, the percentage of mothers who correctly mentioned danger signs of childhood illness did not improve between baseline and follow-up surveys (**Table 42**). Except for fever (mentioned by about 95 percent and 97 percent in the baseline and follow-up surveys, respectively), none of the danger signs of childhood illness were mentioned by more than 45 percent of the mothers. The study arms were different at baseline for the percentage of mothers who mentioned not being able to drink/breastfeed and for the percentage mentioning at least two signs. *Tubaramure's* post-program impact was limited to the percentage of mothers who mentioned rapid breathing as a danger sign of childhood illnesses (4.1 percentage point overall effect, and 8.0 percentage point effect in the TNFP group) (**Table 43**).

At follow-up, only around 10 percent of mothers reported—incorrectly—that sick children younger than 6 months and sick children 6–23 months should be breastfed less (**Table 44**, data collected only at follow-up); around 17 percent wrongly believed that sick children should be provided with less liquid, and around 22 percent thought that they should be given less food. The percentage of mothers who erroneously believed that breast milk, liquid, and food intake should be reduced during convalescence was considerably smaller (below 5 percent for all outcomes).

Tubaramure had a positive post-program effect on mother's knowledge of appropriate feeding for sick children. The intervention increased the percentage of mothers aware of the need to increase breastfeeding for both sick children under 6 months and sick children 6–23 months by around 10 percentage points (**Table 45**). A post-program reduction of about 7 percentage points was found for the percentage who believed sick children should be given fewer liquids. The post-program effect on the percentage of mothers aware that a sick child 6–24 months of age should be provided more solid food was limited to the TNFP group. *Tubaramure* did not have a significant post-program effect on maternal knowledge of how to feed a child recovering from illness.

Virtually all mothers knew that a baby should be breastfed immediately or very soon after birth (baseline and follow-up) and that a baby should be fed colostrum (measured at follow-up only) (**Table 46**). Very few (under 5 percent at both surveys), however, knew that a malnourished mother is capable of producing enough milk to adequately feed her child. When asked about the benefits of exclusive breastfeeding, more than three-quarters of the mothers mentioned benefits related to child health and nutrition at baseline, and this increased to around 89 percent at follow-up. Lactational amenorrhea was largely unknown as a benefit at both baseline and follow-up. At both baseline and follow-up, mothers stated that it was appropriate to stop breastfeeding at around 32 months of age (which meets the WHO recommendation for continued breastfeeding for the first 2 years of life or beyond), but only 10 percent of mothers knew that they could continue breastfeeding when pregnant again. Most mothers (68.6 percent) wrongly believed that if they are unable to breastfeed a child under 6 months of age that the child should be fed cow's or goat's milk; only 22.3 percent suggested breast milk be fed to a child in the mother's absence (data collected only at follow-up).

Tubaramure had a small positive post-program effect on mothers' knowledge that a baby should be fed colostrum (increasing the percentage by 2 to 3 percentage points in all arms), but did not have a clear post-program effect on any of the other breastfeeding knowledge outcomes[43] (**Table 47**).

[43] Note that the significant (simple-difference) effect on the percentage of mothers who believed a child should be fed powdered milk (increased) or baby formula (decreased) might be due to pre-existing differences between groups.

Table 42. Knowledge among mothers of childhood illness danger signs[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3563**	**2568**	**985**	**585**	**998**	**995**	**3559**	**2982**	**989**	**997**	**996**	**577**
Mentioned danger signs of childhood illness												
Cannot drink/breastfeed	22.6*	23.9	21.9	20.7	27.9	19.3	22.1	22.4	24.0	20.6	22.8	20.3
Symptoms intensify	41.1	41.4	42.8	41.0	40.3	40.4	43.3	42.2	46.1	42.4	38.1	49.0
Fever	94.6	94.4	93.7	94.0	95.2	95.1	97.1*	96.8	97.1	96.2	97.2	98.4
Rapid breathing	8.6	8.1	9.4	6.8	7.6	9.7	8.2	8.6	7.1	6.5	12.2	6.2
Difficulty breathing	9.5	10.0	12.0	8.7	8.7	8.4	8.7	9.0	9.7	7.7	9.6	6.9
Bloody stools	12.2	12.1	11.2	11.8	13.2	12.6	10.0	10.1	9.7	12.3	8.2	9.5
Difficulty swallowing	3.3	3.7	4.1	2.1	4.3	2.3	3.7	4.0	4.9	4.1	3.0	2.4
At least two signs	65.1*	66.9	69.1	62.6	67.3	60.2	65.9	65.3	67.9	63.1	64.9	69.0

[a] Values are %.
[b] Sample size at baseline ranged from N = 3562 to 3563 in the full sample; N = 2567 to 2568 in the treatment arms; and N = 984 to 985 in the T24 arm.
[c] Sample size at follow-up did not vary.
* Study arms differed, $p < 0.05$.

Table 43. Knowledge among mothers of childhood illness danger signs: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Mentioned danger signs of childhood illness				
Cannot drink/breastfeed	-1.9 ± 3.7	1.3 ± 4.5	-0.9 ± 4.6	-5.8 ± 4.1
Symptoms intensify	-7.6 ± 7.3	-4.7 ± 8.7	-7.6 ± 9.5	-10.5 ± 8.3
Fever	-0.8 ± 1.0	0.0 ± 1.2	-1.3 ± 1.4	-1.3 ± 1.3
Rapid breathing	4.1 ± 2.0*	1.0 ± 2.5	2.8 ± 2.3	8.0 ± 2.7*
Difficulty breathing	0.8 ± 2.2	-0.6 ± 2.6	0.5 ± 2.9	2.5 ± 3.1
Bloody stools	0.6 ± 3.5	1.4 ± 3.7	3.2 ± 4.5	-2.2 ± 4.9
Difficulty swallowing	0.2 ± 1.6	0.7 ± 2.6	1.8 ± 2.1	-1.5 ± 2.1
At least two signs	-9.6 ± 6.1	-9.5 ± 7.3	-8.3 ± 7.2	-10.6 ± 8.4

[a] Values are double-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 7121 to 7122.
[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.
* Impact estimate significantly different from 0, $p < 0.05$.

Table 44. Knowledge among mothers of feeding practices for sick children[a]

	Follow-up[b]					
	Full sample	Study arm				
		All treated	T24	T18	TNFP	Control
N	**3559**	**2981**	**989**	**997**	**996**	**578**
Feeding a sick child						
Breastfeeding (< 6 months)						
Less	9.4	8.2	4.9	8.7	10.9	15.7
Same	15.6	15.3	17.5	16.0	12.6	16.8
More	74.4	76.0	77.0	74.6	76.3	66.4
Breastfeeding (6–23 months)						
Less	10.5	9.6	7.0	10.5	11.4	14.9
Same	18.3	17.7	20.2	17.1	15.7	21.8
More	70.7	72.2	72.4	71.5	72.9	62.6
Providing liquids (6–23 months)						
Less	16.9	15.9	14.9	15.3	17.4	22.5
Same	18.6	18.5	21.5	17.1	17.0	18.7
More	64.0	65.1	63.3	66.9	65.2	58.1
Providing solid foods (6–23 months)						
Less	22.3	21.4	21.7	20.8	21.7	27.2
Same	22.2	22.1	24.0	22.9	19.3	23.0
More	55.1	56.2	54.1	55.8	58.7	49.3
Feeding a child recovering from illness						
Breastfeeding (< 6 months)						
Less	1.9	1.7	0.3	2.2	2.6	2.8
Same	25.0	25.3	25.8	26.7	23.3	23.7
More	72.8	72.7	73.7	70.6	73.9	73.2
Breastfeeding (6–23 months)						
Less	2.3	2.3	1.6	3.3	1.9	2.4
Same	26.7	26.3	25.7	29.0	24.1	29.1
More	70.6	71.1	72.4	67.2	73.7	68.3
Providing liquids (6–23 months)						
Less	4.5	4.7	4.3	5.1	4.7	3.1
Same	27.5	27.2	26.1	30.5	24.8	29.2
More	67.8	67.8	69.4	63.8	70.3	67.4
Providing solid foods (6–23 months)						
Less	2.8	2.9	2.6	3.5	2.4	2.2
Same	27.0	26.9	26.3	28.9	25.6	27.5
More	70.0	70.0	70.7	67.3	71.9	70.1

[a] Values are %.

[b] Sample size at follow-up ranged from N = 3519 to 3559 in the full sample; N = 2953 to 2981 in the treatment arms; N = 978 to 989 in the T24 arm; N = 987 to 997 in the T18 arm; N = 988 to 996 in the TNFP arm; and N = 566 to 578 in the control arm.

* Study arms differed, $p < 0.05$.

Table 45. Knowledge among mothers of feeding practices for sick children: post-program impact[a]

	Post-program impact[b]			
	T vs. control	**T24 vs. control**	**T18 vs. control**	**TNFP vs. control**
Feeding a sick child				
Breastfeeding (< 6m)				
Less	-7.5 ± 3.3*	-10.8 ± 3.2*	-7.0 ± 3.9	-4.8 ± 4.1
Same	-1.4 ± 2.6	0.7 ± 3.2	-0.8 ± 3.5	-4.2 ± 3.0
More	9.5 ± 3.7*	10.6 ± 4.3*	8.1 ± 4.5	9.9 ± 4.2*
Breastfeeding (6–23 m)				
Less	-5.3 ± 2.9	-7.9 ± 3.2*	-4.4 ± 3.4	-3.5 ± 3.7
Same	-4.2 ± 2.0*	-1.6 ± 2.3	-4.8 ± 3.0	-6.2 ± 2.4*
More	9.7 ± 3.1*	9.8 ± 4.0*	8.9 ± 3.8*	10.3 ± 3.6*
Providing liquids (6–23 m)				
Less	-6.7 ± 3.3*	-7.6 ± 3.9	-7.2 ± 3.8	-5.1 ± 4.4
Same	-0.2 ± 3.0	2.8 ± 3.5	-1.6 ± 3.3	-1.7 ± 3.2
More	7.1 ± 4.3	5.3 ± 5.8	8.8 ± 4.9	7.2 ± 4.9
Providing solid foods (6–23 m)				
Less	-5.8 ± 3.9	-5.5 ± 5.1	-6.3 ± 4.3	-5.4 ± 4.5
Same	-0.9 ± 2.8	1.0 ± 3.2	-0.1 ± 3.2	-3.7 ± 3.2
More	6.9 ± 4.5	4.8 ± 6.4	6.5 ± 5.2	9.4 ± 4.5*
Feeding a child recovering from illness				
Breastfeeding (< 6m)				
Less	-1.1 ± 1.5	-2.5 ± 1.5	-0.6 ± 1.6	-0.2 ± 1.8
Same	1.6 ± 3.4	2.1 ± 4.8	3.0 ± 4.1	-0.4 ± 3.5
More	-0.5 ± 4.1	0.5 ± 5.3	-2.6 ± 4.6	0.7 ± 4.2
Breastfeeding (6–23 m)				
Less	-0.1 ± 0.9	-0.8 ± 0.9	0.9 ± 1.1	-0.5 ± 1.0
Same	-2.8 ± 3.7	-3.4 ± 4.8	-0.1 ± 4.7	-4.9 ± 3.7
More	2.7 ± 4.1	4.0 ± 5.3	-1.1 ± 4.7	5.3 ± 4.1
Providing liquids (6–23 m)				
Less	1.6 ± 1.1	1.1 ± 1.7	2.0 ± 1.6	1.6 ± 1.7
Same	-2.0 ± 3.4	-3.0 ± 4.6	1.3 ± 5.1	-4.3 ± 3.4
More	0.5 ± 3.6	2.0 ± 5.3	-3.6 ± 4.7	3.0 ± 3.7
Providing solid foods (6–23 m)				
Less	0.6 ± 0.9	0.4 ± 1.3	1.3 ± 1.4	0.2 ± 1.1
Same	-0.6 ± 4.4	-1.2 ± 5.2	1.4 ± 5.7	-2.0 ± 4.3
More	-0.1 ± 4.7	0.7 ± 5.8	-2.8 ± 5.6	1.9 ± 4.8

[a] Values are simple-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 3519 to 3559.
* Impact estimate significantly different from 0, $p < 0.05$.

Table 46. Breastfeeding knowledge among mothers[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3562**	**2567**	**985**	**585**	**998**	**995**	**3559**	**2981**	**989**	**997**	**995**	**578**
Knew that:												
Baby should be breastfed immediately or during first hours after birth	94.4	94.7	94.9	94.2	94.7	93.7	96.2	96.4	95.8	96.4	97.2	95.1
Baby should be fed colostrum							98.6*	98.9	99.3	98.7	98.8	96.7
Malnourished mother can produce enough good milk	2.0	1.9	2.2	1.4	1.9	2.3	4.6*	4.4	2.0	5.0	6.1	5.6
Mentioned relation of exclusive breastfeeding to:												
Child health and nutrition	77.3	78.0	78.8	74.8	79.0	75.8	88.8*	90.0	90.8	86.8	92.4	82.9
Appropriate age to stop breastfeeding (months)	31.6 ± 7.3*	31.3 ± 7.4	30.9 ± 7.6	32.3 ± 6.9	31.2 ± 7.3	32.4 ± 7.1	31.8 ± 9.8	31.7 ± 9.6	31.8 ± 8.9	32.3 ± 10.2	31.0 ± 9.4	32.6 ± 10.9
Believe mother can continue breastfeeding when pregnant	–	–	–	–	–	–	10.3	10.5	9.7	9.9	11.8	9.8
If not with child < 6 months, mentioned what to feed:												
Breast milk	–	–	–	–	–	–	22.3	23.4	24.6	22.7	22.7	17.1
Powdered milk	–	–	–	–	–	–	19.7*	20.8	24.4	16.3	21.7	13.8
Baby formula	–	–	–	–	–	–	10.6	9.3	8.5	7.1	12.2	17.6
Cow's or goat's milk	–	–	–	–	–	–	68.6	69.2	70.0	67.6	69.9	65.7
Nothing or does not know	–	–	–	–	–	–	10.2	9.9	7.9	12.5	9.2	12.2

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3507 to 3562 in the full sample; N = 2525 to 2567 in the treatment arms; N = 968 to 985 in the T24 arm; N = 573 to 585 in the T18 arm; N = 984 to 998 in the TNFP arm; and N = 982 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3531 to 3559 in the full sample; N = 2959 to 2981 in the treatment arms; N = 979 to 989 in the T24 arm; N = 992 to 997 in the T18 arm; N = 987 to 995 in the TNFP arm; and N = 572 to 578 in the control arm.
* Study arms differed, p < 0.05.

Table 47. Breastfeeding knowledge among mothers: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	**T24 vs. control**	**T18 vs. control**	**TNFP vs. control**
Knew that:				
Baby should be breastfed immediately or during first hours after birth	0.4 ± 2.6	-0.7 ± 2.8	0.8 ± 3.1	1.1 ± 2.8
Baby should be fed colostrum	2.2 ± 0.7*	2.6 ± 0.6*	2.0 ± 0.8*	2.1 ± 0.8*
Malnourished mother can produce enough good milk	-0.8 ± 1.4	-3.5 ± 1.4	0.5 ± 1.5	1.0 ± 1.6
Mentioned relation of exclusive breastfeeding to:				
Child health and nutrition	5.8 ± 3.9	5.2±4.7	5.3 ± 4.8	6.8 ± 4.8
Appropriate age to stop breastfeeding (months)	0.2 ± 1.1	0.9±1.3	-0.1 ± 1.6	-0.3 ± 1.3
Believed mother can continue breastfeeding when pregnant	0.7 ± 2.5	0.0±2.5	0.1 ± 2.6	2.0 ± 3.6
If not with child < 6 months, mentioned what to feed:				
Breast milk	6.2 ± 3.7	7.5 ± 4.2	5.6 ± 4.0	5.6 ± 4.8
Powdered milk	7.0 ± 2.9*	10.6 ± 4.0*	2.5 ± 4.3	7.9 ± 3.8*
Baby formula	-8.4 ± 4.9	-9.1 ± 5.1	-10.5 ± 5.0*	-5.5 ± 5.5
Cow’s or goat’s milk	3.4 ± 6.1	4.3 ± 6.3	1.9 ± 6.5	4.2 ± 7.2

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points when data from both surveys were available and simple-difference impact estimates ± SE when only follow-up data were available. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 3548 to 7118.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.

* Impact estimate significantly different from 0, $p < 0.05$.

The majority of mothers were aware of the importance of a sufficient quantity of food to prevent malnutrition among children (76.1 percent at baseline and 80.6 percent at follow-up), but a much lower percentage mentioned the importance of dietary quality (around 40 percent at both surveys) or the contribution of illness (approximately 35 percent at baseline and 27 percent at follow-up) (**Table 48**). Knowledge of nutrient-rich foods was limited in both surveys, with fewer than 40 percent of women identifying yellow or orange fruits and vegetables, or animal-source foods, as vitamin A-rich foods. Only green leafy vegetables were mentioned by more than half of mothers (56.4 percent at baseline and 61.3 percent at follow-up). More than two-thirds of mothers were aware that vitamin A deficiency could cause poor immunity in children, whereas only 9 percent were aware that vitamin A deficiency could have vision-related consequences (data collected only in the follow-up survey). For iron-rich foods, mothers primarily identified green leafy vegetables (94.7 percent); fewer than half mentioned meat (45.1 percent), less than 13 percent mentioned CSB, and only 6 percent mentioned special baby foods (data collected only at follow-up). At baseline, around 60 percent of mothers thought that iron deficiency could cause poor immunity, and a similar percentage knew that it could delay development. The percentage of mothers who knew that poor immunity was a consequence of iron deficiency increased to 70.8 percent at follow-up, but the percentage who knew that iron deficiency could delay development did not fundamentally change. Other key consequences of iron deficiency (such as weakness and fatigue) were mentioned by fewer than 30 percent of mothers in both surveys.

The intervention had no clear impact on the percentage of mothers who knew the reasons for child malnutrition (**Table 49**). A strong post-program impact was observed on the percentage of mothers who were able to identify vitamin A-rich foods: the program increased the percentage of mothers mentioning yellow and orange fruits and vegetables (by 13.2 percentage points), green leafy vegetables (by 17.8 percentage points), eggs (by 9.1 percentage points), and liver (by 4.5 percentage points). Moreover, the intervention significantly increased the percentage of mothers who knew that impaired vision was a consequence of vitamin-A deficiency[44] (by 4.7 percentage points). The post-program impact on the percentage of mothers who were able to identify iron-rich foods was limited to CSB[45] (a 9.3 percentage point impact overall, significant in all arms) and green leafy vegetables (only significant in the T24 arm, with a 5.9 percentage point impact). No clear post-program impact was found on knowledge of key consequences of iron deficiency.

[44] Note that this simple-difference effect could be due to pre-existing differences between study arms. The double-difference impact estimates on the other vitamin A-related knowledge outcomes, however, indicate that this program impact is plausible.

[45] Note that this simple-difference effect could be due to pre-existing differences between study arms. Since the consumption of CSB was promoted in the program's BCC, we believe that this program impact is plausible.

Table 48. Malnutrition knowledge among mothers[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3563**	**2568**	**985**	**585**	**998**	**995**	**3559**	**2981**	**989**	**997**	**995**	**578**
Mentioned as reasons for child malnutrition												
Insufficient amount of food	76.1	76.3	74.8	76.9	77.4	75.6	80.6	80.5	77.2	82.6	81.5	80.9
Irregular meals	7.9	7.0	6.4	7.2	7.4	10.3	8.6	8.9	6.3	7.3	13.2	6.8
Diseases	34.9*	36.6	35.8	35.6	38.0	30.5	26.8	26.2	30.7	20.7	27.1	29.8
Early weaning	18.9	18.9	19.4	20.2	17.7	19.0	16.7	17.5	18.3	13.0	21.3	12.5
Lack of affection during feeds	0.9	0.9	0.6	0.5	1.5	0.9	1.4	1.5	1.5	1.5	1.6	0.9
No food variety	39.4	39.2	37.3	39.0	41.2	40.0	42.4	43.4	45.7	42.9	41.7	36.9
Foods perceived as vitamin A rich												
Yellow/orange fruits and vegetables (yellow/orange color)	38.2	36.5	33.9	35.7	39.4	42.8	34.8	35.9	31.5	33.9	42.1	29.5
Green leafy vegetables	56.4	54.3	51.9	55.9	55.8	61.9	61.3*	62.8	60.8	58.0	69.7	53.6
Eggs	10.6	10.8	12.4	9.9	9.6	10.2	16.3*	17.8	17.7	16.4	19.4	8.3
Liver	3.4	3.1	2.9	1.9	4.0	4.1	5.6	6.2	7.6	5.8	5.2	2.8
Breast milk	10.5	10.1	9.6	9.7	10.9	11.6	9.8	9.9	8.7	9.2	11.9	8.8
Cow's milk	12.2	12.2	12.2	8.9	14.0	12.5	11.7	12.3	10.1	11.8	15.0	8.3
Mentioned as consequences of vitamin A deficiency among children												
Vision	–	–	–	–	–	–	9.0*	9.7	8.6	10.5	10.1	5.1
Poor immunity	–	–	–	–	–	–	69.2	70.0	69.5	66.7	73.9	65.1
Foods perceived as iron rich												
Meat	–	–	–	–	–	–	45.1	45.7	50.4	43.0	43.8	41.9
Special baby food	–	–	–	–	–	–	6.2	5.4	5.1	3.3	7.8	10.6
CSB	–	–	–	–	–	–	12.3*	13.8	12.0	10.3	19.0	4.5
Green leafy vegetables	–	–	–	–	–	–	94.7*	95.3	97.3	94.9	93.8	91.3
Mentioned as consequences of iron deficiency among children												
Difficulty in school	1.8	2.0	1.5	2.2	2.3	1.2	1.8	2.0	1.3	1.7	2.9	0.9
Altered development	56.8	57.9	57.4	56.4	59.2	54.0	57.2	57.2	57.4	51.1	63.1	57.4
Slow growth	18.3	18.7	19.3	18.5	18.1	17.5	20.6	21.1	21.6	17.0	24.9	17.8
Poor immunity	59.0	59.1	62.6	57.9	56.4	58.7	70.8	71.2	71.6	71.0	71.2	68.3
Fatigue	13.4*	12.9	12.2	9.7	15.5	14.5	6.5	6.5	4.4	6.2	9.0	6.2
Weakness	27.9	27.5	27.3	23.4	30.3	28.8	21.2	19.7	20.9	16.2	22.0	28.9

[a] Values are %.
[b] Sample size at baseline ranged from N = 3557 to 3563 in the full sample; N = 2563 to 2568 in the treatment arms; N = 983 to 985 in the T24 arm; N = 995 to 998 in the TNFP arm; and N = 994 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3526 to 3559 in the full sample; N = 2956 to 2981 in the treatment arms; N = 976 to 989 in the T24 arm; N = 991 to 997 in the T18 arm; N = 989 to 995 in the TNFP arm; and N = 570 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

Table 49. Malnutrition knowledge among mothers: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Mentioned as reasons for child malnutrition				
Insufficient amount of food	-1.2 ± 2.7	-2.9 ± 3.5	0.6 ± 3.7	-0.9 ± 2.9
Irregular meals	5.5 ± 2.6*	3.3 ± 2.6	3.6 ± 3.1	9.1 ± 3.3*
Diseases	-9.6 ± 5.0	-4.9 ± 5.8	-14.7 ± 5.5	-10.6 ± 5.6
Early weaning	5.5 ± 4.8	5.2 ± 6.2	-0.4 ± 5.3	10.3 ± 5.8*
Lack of affection during feeds	0.6 ± 0.6	0.9 ± 0.6	0.9 ± 0.8	0.1 ± 0.6
No food variety	7.2 ± 4.7	11.7 ± 6.4*	6.2 ± 7.3	3.5 ± 5.4
Foods perceived as vitamin A rich				
Yellow/orange fruits and vegetables	13.2 ± 4.6*	11.2 ± 5.1*	11.4 ± 6.8	16.5 ± 5.2*
Green leafy vegetables	17.8 ± 4.4*	18.0 ± 5.0*	10.5 ± 5.3*	23.1 ± 5.1*
Eggs	9.1 ± 2.4*	7.0 ± 2.9*	8.5 ± 3.3*	11.6 ± 3.0*
Liver	4.5 ± 1.5*	6.1 ± 2.4*	5.1 ± 1.8*	2.5 ± 1.7
Breast milk	2.5 ± 2.8	1.7 ± 3.3	2.1 ± 3.3	3.5 ± 3.9
Cow's milk	4.4 ± 2.9	2.0 ± 3.4	6.8 ± 3.8*	5.1 ± 3.5
Mentioned as consequences of vitamin A deficiency among children				
Vision	4.7 ± 1.4*	3.5 ± 2.2	5.4 ± 1.8*	5.0 ± 1.9*
Poor immunity	4.9 ± 5.4	4.4 ± 6.6	1.6 ± 6.7	8.8 ± 6.1
Foods perceived as iron-rich				
Meat	3.8 ± 6.6	8.5 ± 6.9	1.2 ± 7.7	2.0 ± 7.9
Special baby food	-5.2 ± 4.7	-5.5 ± 4.8	-7.2 ± 4.8	-2.7 ± 5.0
CSB	9.3 ± 1.7*	7.5 ± 2.1*	5.8 ± 2.1*	14.5 ± 2.8*
Green leafy vegetables	4.0 ± 2.8	5.9 ± 2.8*	3.5 ± 2.8	2.4 ± 3.3
Mentioned as consequences of iron-deficiency among children				
Delays studies	0.3 ± 0.8	0.1 ± 1.0	-0.2 ± 1.0	0.9 ± 1.3
Delays development	-2.5 ± 4.8	-2.5 ± 5.5	-8.2 ± 5.5	1.6 ± 6.7
Slow growth	2.5 ± 4.2	1.8 ± 5.2	-2.0 ± 4.4	6.5 ± 5.0
Poor immunity	2.5 ± 5.1	-0.5 ± 6.7	2.8 ± 6.6	5.2 ± 5.3
Fatigue	2.1 ± 1.9	0.3 ± 2.1	4.8 ± 2.8*	1.8 ± 2.7
Weakness	-7.4 ± 5.6	-6.8 ± 7.1	-7.2 ± 6.1	-8.2 ± 6.1

[a] Values are double-difference impact estimates ± SE in percentage points when data from both surveys were available and simple-difference impact estimates ± SE when only follow-up data were available. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 3526 to 7121.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.

* Impact estimate significantly different from 0, $p < 0.05$.

Knowledge related to adequate complementary feeding practices was limited. At baseline, around 62 percent of mothers knew that 6 months is the correct age to introduce foods and liquids other than breast milk (**Table 50**). At follow-up (when questions were asked separately for liquids and foods), 65.0 percent mentioned 6 months as the appropriate age for liquids, and 49.4 percent mentioned 6 months as the appropriate age for the introduction of solid foods. Nearly half of the mothers incorrectly responded that foods should be introduced after 6 months of age. At follow-up, around 27 percent of the mothers erroneously believed that the appropriate food consistency for children 6–8 months of age should be liquid like water (question not asked at baseline). For children 12–23 months, nearly all mothers (99.4 percent) reported that the food should be either thick like a paste or semi-solid like a puree.

When asked about the number of times a child 6–8 months of age should be fed solid (i.e., complementary) foods, the average reported frequency was 2.6 times per day at baseline and 3.1 times per day at follow-up. When the answers were compared to the WHO guidelines of at least twice daily for this age group (WHO 2008), the majority of mothers (93.5 percent at baseline and 98.1 percent at follow-up) knew the minimum frequency; however, when compared with the *Tubaramure* guidelines of thrice daily, 54.7 percent of mothers at baseline and 84 percent at follow-up stated the minimum frequency. Knowledge of the correct feeding frequency for children 12–23 months of age was lower. Mothers reported an average minimum of 2.9 meals daily at baseline and 3.1 meals daily at follow-up. Compared with the WHO guidelines for this age group (at least three meals per day) (WHO 2008), 72.1 percent of mothers at baseline and 88.6 percent at follow-up stated at least the minimum frequency. Knowledge of the *Tubaramure* recommended frequency (at least four meals per day) was considerably lower, with 13.9 percent at baseline and 16.8 percent at follow-up. At follow-up, almost all (97.1 percent) knew that a child 6–8 months of age cannot always eat without help (question not asked at baseline), whereas only 64 percent knew this for children 12–23 months of age.

Tubaramure had positive post-program effects on knowledge regarding the age of introduction of liquids and solid foods other than breast milk[46] (**Table 51**). The program significantly reduced the percentage of mothers who thought that liquids other than breast milk should be introduced before 6 months of age (a reduction of 9 to 14 percentage points, statistically significant in all groups combined and in each group separately) and the percentage who thought that foods should be introduced before 6 months of age (a reduction of 2 to 4 percentage points, statistically significant overall and in the T18 and TNFP groups). Importantly, the program did not significantly increase the percentage of mothers who wrongly believed that complementary foods should be introduced after 6 months of age.

Tubaramure's post-program impact on other feeding-related knowledge was limited. There was a small overall reduction[47] in the percentage of mothers who believed that foods for children 12–23 months of age should be liquid (also significant in the T24 and TNFP groups), and a 5.6 point increase in the percentage of mothers who knew the *Tubaramure*-recommended feeding frequency for children 12–23 months of age (also statistically significant in the T18 group). No clear effect was found on any of the other outcomes related to feeding knowledge.

[46] Note that this simple-difference impact could be due to pre-existing differences between study arms. The groups, however, were balanced with respect to the "age of introduction of food/liquids other than breast milk" variable (i.e., groups were comparable), indicating that this program impact is plausible.

[47] Note that this simple-difference impact could be due to pre-existing differences between study arms.

Table 50. Complementary feeding knowledge among mothers[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	3559	2565	984	585	996	994	3556	2978	988	995	995	578
Age of introduction of food/liquids other than breast milk												
Before 6 months	17.0	17.4	17.3	17.6	17.5	15.9	–	–	–	–	–	–
At 6 months	61.8	61.1	58.8	64.1	61.5	63.7	–	–	–	–	–	–
After 6 months	20.7	21.0	23.2	17.6	20.9	19.9	–	–	–	–	–	–
Does not know	0.5	0.5	0.7	0.7	0.1	0.5	–	–	–	–	–	–
Age of introduction of liquids other than breast milk												
Before 6 months	–	–	–	–	–	–	9.6	7.8	9.7	8.6	5.1	18.7
At 6 months	–	–	–	–	–	–	65.0	65.7	59.7	64.7	72.5	61.4
After 6 months	–	–	–	–	–	–	24.6	25.8	29.8	26.1	21.4	18.5
Does not know	–	–	–	–	–	–	0.8*	0.7	0.7	0.5	1.0	1.4
Age of introduction of foods other than breast milk												
Before 6 months	–	–	–	–	–	–	3.0	2.6	3.2	2.7	1.7	5.4
At 6 months	–	–	–	–	–	–	49.4	49.4	45.0	50.4	52.7	49.3
After 6 months	–	–	–	–	–	–	47.5	48.0	51.6	46.6	45.6	45.2
Does not know	–	–	–	–	–	–	0.1	0.1	0.1	0.3	0.0	0.2
Food consistency for child 6–8 months												
Thick like a paste	–	–	–	–	–	–	1.2	1.2	0.9	1.9	0.9	1.2
Liquid like water	–	–	–	–	–	–	26.5	26.3	24.2	24.8	30.0	27.6
Semi-solid like puree	–	–	–	–	–	–	72.2	72.4	74.9	73.3	69.1	71.2
Food consistency for child 12–23 months												
Thick like a paste	–	–	–	–	–	–	40.2	40.5	47.0	35.6	39.1	38.4
Liquid like water	–	–	–	–	–	–	0.6	0.4	0.5	0.6	0.2	1.7
Semi-solid like puree	–	–	–	–	–	–	59.2	59.0	52.5	63.8	60.7	59.9
Feeding frequency for child 6–8 months												
Number of meals per day	2.6 ± 0.8	2.6 ± 0.8	2.7 ± 0.7	2.6 ± 0.9	2.6 ± 0.8	2.5 ± 0.8	3.1 ± 0.8	3.1 ± 0.8	3.1 ± 0.8	3.1 ± 0.8	3.1 ± 0.7	3.0 ± 0.9
Correct feeding frequency (WHO ≥ 2 meals/day)	93.5	93.6	95.7	93.7	91.6	93.1	98.1	98.3	98.7	97.6	98.8	96.9
Correct feeding frequency (*Tubaramure* ≥ 3 meals/day)	54.7	56.7	63.0	53.8	52.1	49.6	84.0	84.9	85.2	82.8	86.9	79.2
Feeding frequency for child 12–23 months												
Number of meals per day	2.9 ± 1.4	2.9 ± 1.2	3.0 ± 0.7	2.9 ± 1.6	2.8 ± 1.3	2.9 ± 1.7	3.1 ± 1.0	3.1 ± 1.0	3.2 ± 0.7	3.1 ± 1.3	3.1 ± 0.9	3.0 ± 1.0
Correct feeding frequency (WHO ≥ 3 meals/day)	72.1*	72.9	79.8	68.6	68.6	70.2	88.6	89.7	91.4	86.7	91.1	82.8
Correct feeding frequency (*Tubaramure* ≥ 4 meals/day)	13.9	14.2	16.0	12.2	13.7	13.0	16.8*	17.9	20.1	16.9	16.6	11.1
Knew that child cannot always eat without help at:												
6–8 months	–	–	–	–	–	–	97.1	97.4	97.1	97.6	97.5	95.8
12–23 months	–	–	–	–	–	–	64.0	64.6	66.3	64.7	62.8	60.7

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3530 to 3559 in the full sample; N = 2545 to 2656 in the treatment arms; N = 978 to 984 in the T24 arm; N = 581 to 585 in the T18 arm; N = 985 to 996 in the TNFP arm; and N = 982 to 994 in the control arm.
[c] Sample size at follow-up ranged from N = 3509 to 3556 in the full sample; N = 2941 to 2978 in the treatment arms; N = 969 to 988 in the T24 arm; N = 989 to 995 in the T18 arm; N = 982 to 995 in the TNFP arm; and N = 565 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

Table 51. Complementary feeding knowledge among mothers: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Age of introduction of liquids other than breast milk				
Before 6 months	-10.9 ± 4.1*	-9.0 ± 4.3*	-10.1 ± 4.4*	-13.7 ± 4.2*
At 6 months	4.3 ± 5.3	-1.6 ± 5.5	3.4 ± 5.3	11.2 ± 5.9
After 6 months	7.2 ± 2.6*	11.3 ± 3.2*	7.6 ± 3.4*	2.9 ± 3.0
Does not know	-0.6 ± 0.5	-0.7 ± 0.6	-0.9 ± 0.6	-0.4 ± 0.6
Age of introduction of foods other than breast milk				
Before 6 months	-2.8 ± 1.2*	-2.1 ± 1.5	-2.7 ± 1.3*	-3.7 ± 1.3*
At 6 months	0.1 ± 3.7	-4.3 ± 3.8	1.0 ± 5.0	3.4 ± 4.9
After 6 months	2.8 ± 3.8	6.5 ± 4.1	1.5 ± 5.0	0.5 ± 5.2
Does not know	0.0 ± 0.2	-0.1 ± 0.2	0.1 ± 0.2	-0.2 ± 0.2
Food consistency for child 6–8 months				
Thick like a paste	0.0 ± 0.5	-0.3 ± 0.6	0.7 ± 0.8	-0.3 ± 0.5
Liquid like water	-1.3 ± 4.3	-3.4 ± 4.5	-2.8 ± 5.5	2.4 ± 5.2
Semi-solid like puree	1.3 ± 4.3	3.8 ± 4.3	2.2 ± 5.8	-2.1 ± 5.2
Food consistency for child 12–23 months				
Thick like a paste	2.2 ± 7.8	8.6 ± 9.4	-2.8 ± 9.2	0.8 ± 9.0
Liquid like water	-1.3 ± 0.5*	-1.2 ± 0.6*	-1.1 ± 0.6	-1.5 ± 0.5*
Semi-solid like puree	-0.9±8.0	-7.4±9.5	3.9±9.4	0.8±9.2
Feeding frequency for child 6–8 months				
Number of meals per day	0.0 ± 0.1	-0.1 ± 0.1	0.0 ± 0.1	0.1 ± 0.1
Correct feeding frequency (WHO ≥ 2 meals/day)	1.1 ± 1.1	-0.7 ± 1.5	0.1 ± 1.5	3.6 ± 1.5*
Correct feeding frequency (*Tubaramure* ≥ 3 meals/day)	0.1 ± 3.3	-6.5 ± 4.3	0.4 ± 4.5	6.4 ± 3.6*
Feeding frequency for child 12–23 months				
Number of meals per day	0.2 ± 0.1	0.1 ± 0.1	0.2 ± 0.1	0.2 ± 0.1*
Correct feeding frequency (WHO ≥ 3 meals/day)	5.0 ± 4.8	-0.8 ± 4.7	5.6 ± 6.5	10.3 ± 5.2*
Correct feeding frequency (*Tubaramure* ≥ 4 meals/day)	5.6 ± 2.9*	5.7 ± 3.5	6.4 ± 3.3*	4.9 ± 4.1
Knew that child cannot always eat without help at:				
6–8 months	1.6 ± 1.5	1.2 ± 1.9	1.8 ± 1.5	1.7 ± 1.6
12–23 months	4.0 ± 5.7	5.7 ± 6.6	4.1 ± 6.2	2.2 ± 6.4

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points when data from both surveys were available and simple-difference impact estimates ± SE when only follow-up data were available. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 3528 to 7090.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for "Knows correct feeding frequency according to WHO," "Knows correct feeding frequency according to *Tubaramure*," and "Knows that child cannot always eat without help."

* Impact estimate significantly different from 0, $p < 0.05$.

Nearly all mothers knew what is needed to wash hands (more than 99 percent at both surveys mentioned soap; 2.9 percent at baseline and 12.3 percent, at follow-up mentioned ash[48]). However, the importance of washing hands was not well understood at baseline in relation to specific activities, but this generally improved over time (**Table 52**). The majority of mothers correctly stated that it was important to wash one's hands before eating (just less than 95 percent at both surveys), but fewer than 40 percent at baseline mentioned that hands should be washed before feeding a child. This percentage increased to 56.2 percent at follow-up. The percentage who mentioned washing hands after using the toilet increased from 46.2 percent at baseline to 75.9 percent at follow-up, but only 17.0 percent (baseline) and 29.0 percent (follow-up) mentioned washing hands after cleaning a child who had defecated. The average number of specific activities mentioned as key times for handwashing rose from 2.0 at baseline to 2.5 at follow-up.

The most commonly known strategies to prevent worms were washing the child's hands (roughly 69 percent at baseline and 75 percent at follow-up) and careful food preparation (around 36 percent at baseline and 39 percent at follow-up). When asked about how to purify drinking water, 67.0 percent of mothers at baseline and 75.3 percent at follow-up mentioned boiling water. Adding chlorine was largely unknown. At baseline, some of the hygiene knowledge outcomes differed significantly among treatment arms.

The *Tubaramure* intervention had large and positive post-program effects on many of the maternal hygiene knowledge outcomes (**Table 53**). The percentage of mothers who mentioned the following specific handwashing times increased: after toilet use (by 16.6 percentage points for all groups combined, significant in all groups), before feeding a child (by 7.5 percentage points for all groups combined, significant in the TNFP group), and after cleaning a child who defecated (by 8.6 percentage points for all groups combined, significant in the T24 and T18 groups). A 10 point increase was observed in the percentage of mothers naming ash as an appropriate handwashing product (statistically significant overall and in all treatment groups). Overall, there was a positive post-program effect of 0.3 to 0.4 on the average number of key times for handwashing mentioned (in the combined and individual treatment groups). The program also had a positive post-program impact on maternal awareness of several worm-protection methods: washing hands (significant in all groups combined and the T24 group), cutting the child's fingernails (all groups combined and the T18 group), and giving treated water (all groups combined, T24, and TNFP). No significant effect was observed on knowledge of appropriate methods for purifying drinking water.

[48] The use of ash when soap was not available was included in the BCC lessons.

Table 52. Hygiene knowledge among mothers[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3563**	**2568**	**985**	**585**	**998**	**995**	**3560**	**2982**	**989**	**997**	**996**	**578**
Appropriate time for handwashing												
Before eating	94.9	94.4	93.7	93.8	95.5	96.0	93.6	93.7	93.7	92.6	94.7	93.4
After using the toilet	46.2	46.0	43.0	47.5	47.9	46.9	75.9*	78.6	80.5	75.9	79.3	62.2
Before feeding a child	38.3*	39.0	43.7	37.8	35.1	36.6	56.2*	57.7	58.5	54.2	60.5	48.4
After cleaning a child who defecated	17.0	17.3	17.4	18.0	16.9	16.0	29.0*	30.7	32.8	29.8	29.5	20.5
Number of handwashing practices mentioned	2.0 ± 0.9	2.0 ± 0.9	2.0 ± 0.8	2.0 ± 0.9	2.0 ± 0.8	2.0 ± 0.9	2.5 ± 0.9*	2.6 ± 0.9	2.7 ± 0.9	2.5 ± 0.9	2.6 ± 0.9	2.2 ± 0.9
Appropriate handwashing products												
Soap (any)	99.1	99.2	99.2	99.5	99.0	98.9	99.8	99.9	100.0	99.8	99.8	99.7
Ash	2.9	3.0	3.6	2.7	2.6	2.5	12.3*	14.0	14.7	11.5	15.8	3.6
Mud/sand	1.5	1.5	2.4	1.0	0.9	1.5	1.2	1.2	1.1	1.7	0.8	1.2
Appropriate worm-protection methods for children												
Wash hands	68.5	68.7	67.5	66.0	71.5	67.8	74.6	75.9	78.5	72.6	76.7	67.5
Cut fingernails	6.9*	6.7	8.0	3.6	7.2	7.4	4.2	4.6	4.2	5.1	4.4	1.9
Wear pants	2.6*	3.2	2.9	2.7	3.7	1.2	1.1	1.1	1.5	0.8	1.1	1.0
Adequate food preparation	35.5	35.5	32.2	35.0	39.0	35.5	38.6	38.7	39.9	33.4	42.9	37.8
Wear shoes	1.5	1.8	1.9	1.7	1.6	0.9	0.9*	1.0	0.8	0.6	1.5	0.5
Give treated water	11.2	10.9	9.8	12.1	11.1	12.0	12.9	13.6	13.3	10.3	17.1	9.4
Appropriate purification methods for drinking water												
Boiling	67.0	66.9	65.9	65.8	68.5	67.4	75.3	75.4	77.9	71.2	77.0	74.8
Chlorine	3.1	3.3	2.6	3.1	4.0	2.8	2.1	2.0	1.0	1.6	3.3	2.8

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 3559 to 3563 in the full sample; N = 2567 to 2568 in the treatment arms; N = 584 to 585 in the T18 arm; N = 997 to 998 in the TNFP arm; and N = 992 to 995 in the control arm.
[c] Sample size at follow-up ranged from N = 3547 to 3560 in the full sample; N = 2972 to 2982 in the treatment arms; N = 986 to 989 in the T24 arm; N = 993 to 997 in the T18 arm; N = 993 to 996 in the TNFP arm; and N = 575 to 578 in the control arm.
* Study arms differed, $p < 0.05$.

Table 53. Hygiene knowledge among mothers: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Appropriate time for handwashing				
Before eating	1.8 ± 1.2	2.5 ± 1.3*	1.2 ± 1.9	1.6 ± 1.9
After using the toilet	16.6 ± 5.5*	21.2 ± 6.4*	12.4 ± 6.4*	15.2 ± 6.1*
Before feeding a child	7.5 ± 3.7*	3.4 ± 4.5	4.4 ± 5.0	13.8 ± 4.3*
After cleaning a child who defecated	8.6 ± 3.9*	10.6 ± 4.4*	7.0 ± 4.0*	7.7 ± 6.0
Average number of handwashing practices mentioned	0.3 ± 0.1*	0.4 ± 0.1*	0.3 ± 0.1*	0.4 ± 0.1*
Appropriate handwashing products				
Soap (any)	-0.2 ± 0.6	-0.1 ± 0.7	-0.5 ± 0.7	-0.1 ± 0.7
Ash	10.0 ± 2.1*	10.0 ± 3.2*	7.6 ± 2.4*	11.8 ± 3.5*
Mud/sand	-0.1 ± 0.9	-1.2 ± 1.5	0.9 ± 1.3	0.1 ± 0.8
Appropriate worm-protection methods for children				
Wash hands	8.3 ± 3.7*	11.7 ± 4.2*	6.9 ± 5.5	5.9 ± 5.4
Cut fingernails	3.3 ± 1.9*	1.7 ± 2.1	6.6 ± 3.1*	2.6 ± 2.6
Wear pants	-1.8 ± 0.9	-1.2 ± 1.0	-1.7 ± 1.2	-2.4 ± 1.3
Adequate food preparation	1.9 ± 5.9	5.7 ± 8.1	-3.5 ± 6.3	2.0 ± 7.1
Wear shoes	-0.4 ± 0.6	-0.8 ± 0.9	-0.8 ± 0.7	0.3 ± 0.8
Give treated water	5.3 ± 2.2*	5.9 ± 3.1*	0.6 ± 3.2	8.3 ± 3.4*
Appropriate purification methods for drinking water				
Boiling	1.1 ± 3.6	4.3 ± 4.9	-3.0 ± 5.3	0.9 ± 3.9
Chlorine	-1.1 ± 1.6	-1.5 ± 1.8	-1.3 ± 1.6	-0.5 ± 2.3

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 7109 to 7123. One-sided tests were used for all indicators.
[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.
* Impact estimate significantly different from 0, $p < 0.05$.

9.2 Maternal Diet and Nutritional Status

The average dietary diversity score among mothers at follow-up was 4.6 (data not collected at baseline) (**Table 54**). Starchy staples, legumes and nuts, dark green leafy vegetables, vitamin A-rich fruits and vegetables, and other fruits and vegetables were consumed by at least two-thirds of mothers in the past 24 hours. Consumption of other food groups was much less common: only about one-third of mothers had consumed meat and fish, and fewer than 3 percent reported consumption of organ meat, eggs, and dairy products.

Tubaramure had a small and positive post-program impact[49] on dietary diversity (an increase ranging from 0.4 to 0.5 food groups, statistically significant for all groups combined and in all groups separately) (**Table 55**). The largest post-program impact was found in the increased consumption of other fruits and vegetables (16.7 percentage points overall, also significant in all groups separately), legumes and nuts (8.9 percentage points overall, also significant in the TNFP group), and dark green leafy vegetables (6.6 percentage points, also significant in the T18 and TNFP arms). In addition, post-program effects of around 1 percentage point were found for organ meat, eggs, and milk and milk products.

Table 54. Maternal consumption and dietary diversity[a]

	Follow-up[b]					
		Study arm				
	Full sample	All treated	T24	T18	TNFP	Control
N	**3493**	**2920**	**958**	**978**	**984**	**573**
Dietary diversity (range 0 to 9)	4.6 ± 1.1*	4.7 ± 1.1	4.6 ± 1.1	4.6 ± 1.2	4.7 ± 1.1	4.2 ± 1.2
Maternal consumption in previous 24 hours						
Starchy staples	94.2	94.4	93.5	94.1	95.6	93.0
Dark green leafy vegetables	88.7*	89.8	87.8	92.0	89.5	83.2
Other vitamin A-rich fruits and vegetables	89.1	89.1	87.9	89.3	90.1	88.8
Other fruits and vegetables	69.0*	71.7	76.7	67.7	70.9	55.1
Organ meat	1.0	1.2	1.1	0.8	1.5	0.3
Meat and fish	32.8	33.5	30.4	35.9	34.2	29.3
Eggs	1.5	1.7	2.2	1.3	1.6	0.5
Legumes and nuts	79.4	80.8	79.4	78.3	84.6	71.9
Milk and milk products	2.5	2.7	1.7	3.3	3.3	1.2

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at follow-up ranged from N = 18 to 3493 in the full sample; N = 18 to 2920 in the treatment arms; N = 8 to 958 in the T24 arm; N = 3 to 978 in the T18 arm; N = 7 to 984 in the TNFP arm; and N = 0 to 573 in the control arm.
* Study arms differed, $p < 0.05$.

[49] We cannot exclude the possibility that this simple-difference effect was due to differences between groups that existed before the program. The impact on mothers' dietary diversity, however, was consistent with the post-program (double-difference) impact on household dietary diversity (see Table 44).

Table 55. Maternal consumption and dietary diversity: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Dietary diversity (range 0 to 9)	0.4 ± 0.1*	0.4 ± 0.1*	0.4 ± 0.1*	0.5 ± 0.1*
Maternal consumption in previous 24 hours				
Starchy staples	1.4 ± 2.8	0.5 ± 3.0	1.1 ± 3.1	2.6 ± 2.8
Dark green leafy vegetables	6.6 ± 2.7*	4.5 ± 3.2	8.8 ± 2.8*	6.3 ± 3.4*
Other vitamin A-rich fruits and vegetables	0.3 ± 2.3	-0.9 ± 3.0	0.4 ± 2.9	1.3 ± 2.6
Other fruits and vegetables	16.7 ± 5.0*	21.6 ± 5.6*	12.6 ± 6.5*	15.8 ± 5.7*
Organ meat	0.8 ± 0.4*	0.8 ± 0.7	0.5 ± 0.5	1.2 ± 0.6*
Meat and fish	4.2 ± 5.5	1.1 ± 5.5	6.6 ± 6.1	4.9 ± 7.1
Eggs	1.2 ± 0.4*	1.7 ± 0.7*	0.8 ± 0.6	1.1 ± 0.6*
Legumes and nuts	8.9 ± 3.9*	7.5 ± 4.6	6.4 ± 4.2	12.7 ± 4.6*
Milk and milk products	1.5 ± 0.7*	0.4 ± 0.7	2.0 ± 0.8*	2.0 ± 1.4

[a] Values are simple-difference impact estimates ± SE in variable mean or percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 3481 to 3493. One-sided tests were conducted for all outcomes.
[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.
* Impact estimate significantly different from 0, $p < 0.05$.

On average, mothers weighed 51 kg and were 156 cm tall in both surveys. About 3 percent of the mothers were of short stature (i.e., less than 145 cm tall) at both baseline and follow-up. Short stature increases the odds of complications during childbirth. The vast majority of non-pregnant mothers (roughly 81 percent at both baseline and follow-up) had a normal BMI[50] (**Table 56**). Those that fell outside of the normal range were more likely to be underweight (14.2 percent at baseline and 12.8 percent at follow-up) than overweight (4.8 percent at baseline and 5.7 percent at follow-up). Mean Hb dropped from 12.5 g/dL at baseline to 12.0 g/dL at follow-up. This drop in Hb was reflected in a steep increase in the prevalence of anemia, from 30.0 percent of the mothers at baseline to more than 42 percent at follow-up. Severe anemia was virtually absent in both surveys.

No significant post-program impact of *Tubaramure* was observed for maternal weight, but the program had a significant post-program impact on increasing the number of women with a normal BMI (an increase of 6.0 percentage points overall, significant in the T24 and T18 groups) (**Table 57**). This increase in normal weight was largely due to a drop in the number of women with a low BMI, but this effect was only significant in the T24 group. The intervention did not have a post-program protective effect on Hb levels or (severe) anemia.

[50] All non-pregnant women were included. Women who had recently given birth may thus have biased the mean BMI estimate upward.

Table 56. Anthropometric status of mothers[a]

	Baseline[b]						Follow-up[c]					
	Full sample	Study arm					Full sample	Study arm				
		All treated	T24	T18	TNFP	Control		All treated	T24	T18	TNFP	Control
N	3482	2510	967	571	973	972	3500	2940	976	983	981	560
Weight (kg)[d]	50.6 ± 7.0	50.6 ± 7.1	50.6 ± 7.1	50.7 ± 6.8	50.6 ± 7.2	50.4 ± 6.7	51.0 ± 6.6	51.0 ± 6.6	50.8 ± 6.2	51.2 ± 6.5	51.1 ± 7.0	50.8±7.0
Height (cm)	155.8 ± 6.2	156.0 ± 6.2	156.5 ± 5.7	156.2 ± 6.0	155.3 ± 6.8	155.4 ± 6.2	155.8 ± 5.9	156.0 ± 5.8	156.4 ± 5.8	156.0 ± 5.8	155.5 ± 5.9	155.2±6.0
% shorter than 145 cm	3.1*	3.1	1.9	2.6	4.6	3.2	3.1	2.8	2.0	3.0	3.5	4.3
BMI												
Mean BMI[d]	20.9 ± 3.6	20.9 ± 3.5	20.7 ± 2.6	20.8 ± 2.4	21.1 ± 4.5	21.0 ± 3.9	21.0 ± 2.4	21.0 ± 2.4	20.7 ± 2.2	21.0 ± 2.4	21.2 ± 2.6	21.0±2.5
Underweight (BMI < 18.5)	14.2	15.1	15.8	16.0	13.9	11.9	12.8	12.7	12.6	12.9	12.6	13.6
Normal (BMI 18.5–24.9)	81.0	80.2	80.3	79.5	80.4	83.1	81.4	81.9	84.0	81.9	79.6	79.1
Overweight (BMI ≥ 25)	4.8	4.7	3.9	4.6	5.7	5.0	5.7	5.4	3.4	5.2	7.8	7.3
Hemoglobin/anemia												
Mean Hb (g/dL)	12.5 ± 1.6*	12.6 ± 1.6	12.8 ± 1.5	12.2 ± 1.7	12.7 ± 1.6	12.2 ± 1.6	12.0 ± 1.7	12.1 ± 1.7	12.2 ± 1.7	12.0 ± 1.7	12.1 ± 1.6	11.8±1.7
% anemic[e]	30.0*	28.2	24.0	37.5	27.0	34.8	42.2	41.2	39.1	44.0	40.6	47.1
% severely anemic[f]	0.6	0.6	0.2	0.9	0.7	0.8	0.9	0.8	0.9	1.0	0.4	1.3

[a] Values are % or mean ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 2746 to 3482 in the full sample; N = 1987 to 2510 in the treatment arms; N = 778 to 967 in the T24 arm; N = 438 to 571 in the T18 arm; N = 771 to 973 in the TNFP arm; and N = 759 to 972 in the control arm.
[c] Sample size at follow-up ranged from N = 2883 to 3500 in the full sample; N = 2433 to 2940 in the treatment arms; N = 832 to 976 in the T24 arm; N = 806 to 983 in the T18 arm; N = 795 to 981 in the TNFP arm; and N = 450 to 560 in the control arm.
[d] Weight and BMI are reported only for women who reported not being pregnant at the time of the interview.
[e] The cutoff for anemia was 12 g/dL for non-pregnant women and 11 g/dL for pregnant women.
[f] The cutoff for severe anemia was 8 g/dL for non-pregnant women and 7 g/dL for pregnant women.
* Study arms differed, $p < 0.05$.

Table 57. Anthropometric status of mothers: post-program impact[a]

	Post-program impact[b]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Weight (kg)[c]	0.0 ± 0.5	-0.1 ± 0.6	-0.1 ± 0.6	0.2 ± 0.5
BMI				
Mean BMI[c]	0.0 ± 0.2	0.0 ± 0.2	0.0 ± 0.3	0.0 ± 0.3
Underweight (BMI < 18.5)	-4.2 ± 2.5	-5.0 ± 2.8*	-4.4 ± 3.0	-3.1 ± 3.1
Normal (BMI 18.5–24.9)	6.0 ± 2.0*	8.0 ± 2.4*	6.5 ± 2.3*	3.6 ± 2.6
Overweight (BMI ≥ 25)	-1.9 ± 1.7	-3.0 ± 1.7	-2.1 ± 2.0	-0.5 ± 2.0
Hemoglobin/anemia				
Mean Hb (g/dL)	0.0 ± 0.1	-0.1 ± 0.1	0.2 ± 0.2	-0.1 ± 0.1
% anemic[d]	-0.1 ± 4.1	2.8 ± 4.2	-5.7 ± 5.1	1.0 ± 4.6
% severely anemic[e]	-0.3 ± 0.6	0.3 ± 0.7	-0.3 ± 0.8	-0.8 ± 0.7

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 5591 to 5622. One-sided tests were used for maternal weight and for maternal hemoglobin and anemia variables: maternal underweight, adjusted values, maternal anemia, and maternal severe anemia. The anemia and hemoglobin estimates further controlled for altitude, child age, household size, household dependency ratio, mother's age, mother's education, household head education, and household head occupation at baseline and follow-up, and for *colline*-fixed effects.
[c] Weight and BMI are reported only for women who reported not being pregnant at the time of the interview.
[d] The cutoff for anemia was 12 g/dL for non-pregnant women and 11 g/dL for pregnant women.
[e] The cutoff for severe anemia was 8 g/dL for non-pregnant women and 7 g/dL for pregnant women.
* Impact estimate significantly different from 0, $p < 0.05$.

10. Results: Child Post-Program Impact

10.1 Child Health

Utilization and documentation of preventive health services for children 24–41 months of age modestly improved between baseline and follow-up (**Table 58**). Roughly three-fifths of caregivers reportedly had a vaccination card for their children where vaccinations, vitamin A supplementation, height or length, and weight could be recorded (58.4 percent at baseline and 62.7 percent at follow-up). The percentage who had the card available at the interview increased from 40.1 percent at baseline to 50.1 percent at follow-up. At follow-up, around 5 percent of children had attended growth monitoring in the past month.[51] Nearly half of the children had received vitamin A in the past 6 months at baseline (as reported by the mother or recorded on the vaccination card), and 67.9 percent had received it in the past 6 months at follow-up. Among the group of children with a vaccination card available, around 83 percent and 89 percent of children were fully vaccinated for their age at baseline and follow-up, respectively.[52] There were no differences between the study arms at baseline.

Tubaramure did not have a clear post-program impact on any of the preventive health care practices received by children 24-41 months of age (**Table 59**).

Illness was common among the children, with roughly one-third of them experiencing at least one symptom of illness in the past 2 weeks at both baseline and follow-up (**Table 60**). The most commonly reported symptoms were fever (25.1 percent at baseline and 21.5 percent at follow-up), loss of appetite (25.1 percent at baseline and 21.6 percent at follow-up), and coughing (25.7 percent at baseline and 21.4 percent at follow-up). Nearly one in six children at baseline (15.9 percent) and one in nine children at follow-up (11.3 percent) had experienced severe diarrhea (defined as diarrhea with at least one of the following: fever, six or more loose stools per day, vomiting, not wanting to eat or drink, blood in the stool, or parents' belief that the illness is not getting better) in the past 2 weeks, and between 1 percent and 3 percent of children at both surveys had experienced severe respiratory problems[53] (cough with difficulty breathing or fast breathing due to chest problems) in the past 2 weeks. None of the illness outcomes were different between study arms at baseline.

Tubaramure had consistent, large and positive post-program effects on child morbidity in the T24 and TNFP study arms (**Table 61**). The program reduced the percentage of children with morbidity symptoms by 13 to 19 percentage points (significant for all treatment groups combined, and in the T24 and TNFP groups, when compared with the control group). Similarly, there was an impact on the prevalence of the most common symptoms of illness (such as fever, loss of appetite, and cough; significant for all treatment groups combined, and in the T24 and TNFP groups) and for less common (but important) health problems, such as difficulties with drinking, difficulty breathing, convulsions, and watery diarrhea (significant for all treatment groups combined, and in the T24 and TNFP treatment arms). A significant program effect was found in all treatment groups for watery diarrhea and non-watery non-bloody

[51] The recall period at baseline was 2 months, and the sample mean was 10.6 percent. If the *Tubaramure* recommendation (growth monitoring every 3 months) was followed, approximately two-thirds of respondents at baseline and one-third at follow-up would be expected to report having taken their child to the clinic for growth monitoring. The percentages at both surveys were considerably lower, indicating that growth monitoring attendance was lower than recommended.

[52] Note that Burundi's vaccination schedule has the last recommended vaccination at the age of 9 months. Since children in this survey were 24–41 months of age, we could not assess with these data whether children received the immunizations on time. Note that the 2012 report found that among children 0–23 months of age with a vaccination card, only 63.2 percent were fully vaccinated for their age.

[53] Severe respiratory problems might be indicative of pneumonia. Both severe diarrhea and severe respiratory problems require immediate attention from a trained health professional.

diarrhea. The program also had significant, protective effects on severe diarrhea (9 to 12 percentage points, significant in all treatment arms) and the prevalence of potential pneumonia (2.5 percentage points, effect limited to the T24 treatment group).

Roughly 67 percent of children at baseline who had a fever in the past 2 weeks received fever-reducing medication, and this increased to around 77 percent at follow-up (**Table 62)**. Only around 40 percent of those with diarrhea at both surveys, however, had received ORS. At both time points, around 30 percent of mothers who had a child with diarrhea reported reducing the child's liquid intake. This was considerably higher than the percentage of mothers who believed that liquid intake should be reduced in sick children[54] (**Table 44**). The percentage of mothers with a severely ill child who sought care from a trained professional increased from 67.1 percent at baseline to 72.3 percent at follow-up. Very few (< 1 percent at baseline and < 6 percent at follow-up) sought care from someone other than a trained professional. None of these outcomes was significantly different between arms at baseline.

The program did not have a significant post-program impact on the treatment of illness or curative care seeking practices (**Table 63**).

[54] Note that knowledge questions regarding feeding of a sick child were asked with respect to a child 6–23 months of age. See section 9.1.

Table 58. Preventive health care practices[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3561**	**2567**	**984**	**585**	**998**	**994**	**3560**	**2982**	**989**	**997**	**996**	**578**
Vaccination card												
Reported having a vaccination card	58.4	58.9	62.4	54.7	57.8	57.4	62.7	61.9	63.9	56.3	65.5	66.8
Showed a vaccination card	40.1	40.2	44.5	40.0	36.0	40.1	50.1	49.5	51.8	43.0	53.8	53.3
Growth monitoring												
Attended growth monitoring in the past 2 months (baseline) or in the past month (follow-up)	10.6	12.1	13.0	9.6	12.6	6.8	5.4	5.6	5.9	6.0	5.0	4.2
Vitamins and vaccinations												
Received vitamin A in past 6 months	49.2	49.5	49.7	52.3	47.6	48.6	67.9	68.8	70.7	66.0	69.6	63.7
Fully vaccinated for age	83.1	82.9	88.3	76.9	80.2	83.4	88.5	89.2	89.6	87.4	90.3	85.1

[a] Values are %.

[b] Sample size at baseline ranged from N = 253 to 3561 in the full sample; N = 203 to 2567 in the treatment arms; N = 85 to 984 in the T24 arm; N = 31 to 585 in the T18 arm; N = 87 to 998 in the TNFP arm; and N = 50 to 994 in the control arm.

[c] Sample size at follow-up ranged from N = 116 to 3560 in the full sample; N = 99 to 2982 in the treatment arms; N = 33 to 989 in the T24 arm; N = 38 to 997 in the T18 arm; N = 28 to 996 in the TNFP arm; and N = 17 to 578 in the control arm.

* There were no statistical differences among study arms, $p < 0.05$.

Table 59. Preventive health care practices: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Vaccination card				
Reported having a vaccination card	-5.2 ± 4.3	-7.5 ± 5.3	-6.9 ± 4.8	-1.6 ± 6.5
Showed a vaccination card	-2.4 ± 3.9	-5.1 ± 5.1	-9.1 ± 4.8	5.2 ± 5.4
Growth monitoring[d]				
Attended growth monitoring in the last month	1.5 ± 1.1	1.7 ± 1.4	1.9 ± 1.7	0.9 ± 1.4
Vitamins and vaccinations				
Received vitamin A in past 6 months	4.5 ± 6.1	6.5 ± 8.0	-1.8 ± 6.5	7.3 ± 6.8
Fully vaccinated for age	5.5 ± 3.9	1.7 ± 4.5	7.1 ± 6.8	8.6 ± 5.1

[a] Values are double-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 369 to 7121.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.

[d] Simple-difference impact estimates ± SE using follow-up data.

* Impact estimate significantly different from 0, $p < 0.05$.

Table 60. Child health and prevalence of morbidity symptoms[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3562**	**2568**	**985**	**585**	**998**	**994**	**3560**	**2982**	**989**	**997**	**996**	**578**
Symptoms experienced during past 2 weeks												
Any illness	37.3	38.4	42.1	34.1	37.3	34.3	32.0	30.5	29.0	34.5	28.0	39.7
Fever	25.1	24.4	24.4	21.7	26.0	26.8	21.5	19.7	16.1	24.2	18.8	30.5
Convulsions	4.9	4.5	4.6	3.6	4.8	6.1	4.3*	3.5	2.5	5.6	2.3	8.3
Problems drinking	13.2	13.2	14.9	9.4	13.8	13.0	12.2	10.9	10.3	13.5	9.0	18.7
Loss of appetite	25.1	25.4	26.9	22.4	25.6	24.5	21.6	20.0	18.6	21.7	19.7	29.6
Vomited all (s)he ate/drank	12.8	12.2	13.1	10.4	12.2	14.4	10.6	9.6	8.7	10.4	9.6	15.6
Cough	25.7	27.1	30.9	21.9	26.5	22.1	21.4	20.4	19.5	23.2	18.5	26.5
Difficulty breathing	17.0	16.9	18.0	14.2	17.5	17.1	11.9	10.7	11.4	10.6	10.0	18.0
Bloody diarrhea	3.0	2.7	3.2	2.4	2.4	3.9	1.6	1.3	1.0	1.6	1.3	3.0
Watery diarrhea	13.6	14.0	16.0	12.0	13.2	12.8	9.8	8.2	8.5	9.1	6.9	18.0
Other type of diarrhea	4.7	5.0	5.0	3.8	5.6	3.9	4.0	3.3	4.1	2.7	3.2	7.1
Suffered from potentially severe illness requiring immediate medical attention during past 2 weeks												
Severe diarrhea[d]	15.9	16.3	18.2	14.5	15.5	15.0	11.3	9.7	11.0	9.9	8.2	19.8
Severe respiratory problems[e]	2.7	2.8	3.1	2.2	2.8	2.5	1.9	1.7	1.0	1.9	2.3	2.9

[a] Values are %.

[b] Sample size at baseline ranged from N = 3522 to 3562 in the full sample; N = 2539 to 2568 in the treatment arms; N = 967 to 985 in the T24 arm; N = 581 to 585 in the T18 arm; N = 991 to 998 in the TNFP arm; and N = 983 to 994 in the control arm.

[c] Sample size at follow-up ranged from N = 3529 to 3560 in the full sample; N = 2952 to 2982 in the treatment arms; N = 977 to 989 in the T24 arm; N = 987 to 997 in the T18 arm; N = 988 to 996 in the TNFP arm; and N = 575 to 578 in the control arm.

[d] Diarrhea with at least one of the following: fever, six or more loose stools per day, vomiting, not wanting to eat or drink, blood in the stool, or parents' belief that the illness was not improving.

[e] Cough in the past 2 weeks along with difficulty breathing or fast breathing due to chest problems.

* Study arms differed, $p < 0.05$.

Table 61. Child health and prevalence of morbidity symptoms: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Symptoms experienced during the past 2 weeks				
Any illness	-13.5 ± 5.4*	-18.9 ± 6.7*	-4.5 ± 8.4	-14.7 ± 6.5*
Fever	-8.5 ± 4.3*	-11.8 ± 4.7*	-1.0 ± 7.2	-10.7 ± 4.8*
Convulsions	-3.2 ± 1.6*	-4.2 ± 1.5*	0.0 ± 2.9	-4.6 ± 1.7*
Problems drinking	-7.8 ± 2.7*	-10.3 ± 3.2*	-1.2 ± 4.3	-10.2 ± 3.2*
Loss of appetite	-10.3 ± 3.9*	-13.3 ± 4.6*	-5.6 ± 6.0	-10.9 ± 5.4*
Vomited all (s)he ate/drank	-3.5 ± 3.3	-5.3 ± 3.8	-1.0 ± 4.1	-3.6 ± 3.6
Cough	-11.1 ± 4.2*	-16.1 ± 5.6*	-2.7 ± 6.2	-12.4 ± 5.4*
Difficulty breathing	-6.9 ± 2.3*	-7.4 ± 3.1*	-4.3 ± 3.4	-8.3 ± 3.3*
Bloody diarrhea	-0.3 ± 1.0	-1.1 ± 1.3	0.3 ± 1.1	0.0 ± 1.0
Watery diarrhea	-10.7 ± 3.1*	-12.4 ± 4.2*	-7.8 ± 3.5*	-11.3 ± 3.4*
Other type of diarrhea	-4.6 ± 2.2*	-4.1 ± 2.3*	-4.1 ± 2.4*	-5.4 ± 2.5*
Suffered from potentially severe illness requiring immediate medical attention during past 2 weeks				
Severe diarrhea[d]	-11.0 ± 3.8*	-11.8 ± 4.9*	-9.0 ± 4.4*	-11.8 ± 4.2*
Severe respiratory problems[e]	-1.5 ± 1.4	-2.5 ± 1.4*	-0.7 ± 1.7	-0.9 ± 1.6

[a] Values are double-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 7051 to 7122.

[c] One-sided tests were used for all child morbidity symptoms, severe diarrhea, and potential pneumonia. The *a priori* hypothesis was a negative effect for all indicators.

[d] Diarrhea with at least one of the following: fever, six or more loose stools per day, vomiting, not wanting to eat or drink, blood in the stool, or parents' belief that the illness was not improving.

[e] Cough in the past 2 weeks along with difficulty breathing or fast breathing due to chest problems.

* Impact estimate significantly different from 0, $p < 0.05$.

Table 62. Treatment of illness and curative care seeking practices[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**883**	**616**	**232**	**128**	**256**	**267**	**747**	**575**	**155**	**244**	**176**	**172**
Fever												
Received medication for fever	67.0	66.2	60.3	74.2	67.6	68.9	77.0	76.2	71.0	80.3	75.0	79.7
Diarrhea:												
Received ORS	40.2	42.6	48.6	37.2	38.6	33.1	39.8	38.1	40.8	34.7	38.8	43.9
Amount of liquids given												
Nothing or less than normal	34.2	33.7	32.6	37.2	33.1	35.4	27.6	28.1	24.3	23.5	38.1	26.5
Same as normal	14.5	13.9	19.3	9.3	10.2	16.3	11.6	11.6	11.7	14.3	8.3	11.5
More than normal	51.3	52.4	48.1	53.5	56.7	48.3	60.8	60.4	64.1	62.2	53.6	61.9
Children with symptoms requiring immediate medical attention (severe diarrhea or severe respiratory problems) who:												
Sought care/advice from trained professional	67.1	66.5	61.7	60.5	75.5	68.9	72.3	70.4	67.9	72.0	71.6	77.3
Sought care/advice from somebody else	0.2	0.2	0.0	1.2	0.0	0.0	5.3	5.5	2.8	9.3	4.2	5.0
Did not seek care	32.7	33.3	38.3	38.3	24.5	31.1	22.3	24.1	29.4	18.7	24.2	17.6

[a] Values are %.

[b] Sample size at baseline ranged from N = 560 to 883 in the full sample; N = 412 to 616 in the treatment arms; N = 180 to 232 in the T24 arm; N = 81 to 128 in the T18 arm; N = 151 to 256 in the TNFP arm; and N = 147 to 267 in the control arm.

[c] Sample size at follow-up ranged from N = 398 to 747 in the full sample; N = 285 to 575 in the treatment arms; N = 103 to 155 in the T24 arm; N = 98 to 244 in the T18 arm; N = 84 to 176 in the TNFP arm; and N = 113 to 172 in the control arm.

* There were no statistical differences among study arms, $p < 0.05$.

Table 63. Treatment of illness and curative care seeking practices: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Fever				
Received medication for fever	-2.1 ± 8.2	-1.3 ± 9.8	-5.2 ± 8.5	-0.4 ± 8.7
Diarrhea				
Received ORS	-16.6 ± 11.0	-24.2 ± 12.5	-11.2 ± 13.9	-12.5 ± 11.6
Amount of liquids given				
Nothing or less than normal	4.9 ± 6.2	2.1 ± 6.6	0.0 ± 8.6	11.8 ± 8.9
Same as normal	3.2 ± 5.7	0.6 ± 7.3	7.5 ± 6.7	2.7 ± 6.2
More than normal	-8.1 ± 5.4	-2.8 ± 7.1	-7.5 ± 7.4	-14.5 ± 6.8*
Children with symptoms requiring immediate medical attention (severe diarrhea or respiratory problems) who:				
Sought care/advice from trained professional	-5.6 ± 7.0	-3.3 ± 8.1	1.3 ± 9.1	-13.1 ± 8.2
Sought care/advice from somebody else	-0.4 ± 2.4	-3.4 ± 2.5	3.8 ± 2.9	-0.4 ± 3.1
Did not seek care	6.0 ± 7.7	6.7 ± 8.3	-5.2 ± 9.8	13.5 ± 8.9

[a] Values are double-difference impact estimates ± SE in percentage points. All estimates controlled for *colline*-level fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 969 to 1630.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for treatment of fever with medication and treatment of diarrhea with ORS.

* Impact estimate significantly different from 0, $p < 0.05$.

10.2 Child Development

At baseline, the mean highest attained motor milestone (see **Appendix 3** for the full list of motor milestones) among the full sample of children 24–41 months of age was jumping with both feet up at the same time, and there was no statistically significant difference across the groups (**Table 64**). At follow-up, the highest average attained motor milestone increased by about one milestone, to standing on one foot for several seconds. The largest increase over time was seen among children 24–29 months of age (where the highest average motor milestone attained across the full sample of children in this age group changed from walking forward in a straight line for 10 steps to jumping on both feet). Program impact was limited to this age group, where a statistically significant effect of 0.5 motor milestones was found in the T24 arm compared with the control arm (**Table 65**).

The highest average attained language milestone (see **Appendix 4** for the full list of language milestones) for the full sample of children 24–41 months of age at baseline was the ability to say 20 or more words, and the groups were balanced at baseline (**Table 64**). Language development appeared to improve for all age groups between 2010 and 2014. Among children in the two older age groups, the differences between baseline and follow-up were not significantly greater for the treatment arms when compared with the control groups. However, *Tubaramure* significantly improved language development among children 24–29 months of age across all treatment groups combined (a 0.9 point effect) and in the T24 and T18 groups (a 1.0 and 1.5 point effect, respectively) (**Table 65**).

Table 64. Motor and language milestones attained[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	**3545**	**2557**	**982**	**583**	**993**	**988**	**3530**	**2959**	**979**	**990**	**990**	**571**
Highest motor milestone attained												
All (24–41 months)	24.1 ± 3.4	24.2 ± 3.3	24.1 ± 3.3	24.0 ± 3.5	24.3 ± 3.2	24.0 ± 3.5	25.0 ± 3.1	25.0 ± 3.2	25.1 ± 3.1	25.0 ± 3.2	25.0 ± 3.2	24.5 ± 3.0
24–29 months	22.9 ± 3.1	23.0 ± 3.0	22.9 ± 2.9	22.7 ± 3.3	23.2 ± 3.0	22.9 ± 3.4	23.9 ± 3.1	23.9 ± 3.1	24.1 ± 3.0	23.8 ± 3.2	23.8 ± 3.1	23.6 ± 2.8
30–35 months	24.1 ± 3.4	24.2 ± 3.3	24.3 ± 3.3	24.0 ± 3.5	24.3 ± 3.1	23.8 ± 3.6	24.9 ± 3.1	25.0 ± 3.2	25.1 ± 3.1	25.1 ± 3.3	24.9 ± 3.1	24.2 ± 2.9
36–41 months	25.4 ± 3.2	25.4 ± 3.1	25.3 ± 3.3	25.4 ± 3.0	25.5 ± 3.0	25.3 ± 3.3	26.0 ± 2.9	26.1 ± 2.9	26.1 ± 2.8	25.9 ± 3.0	26.2 ± 2.9	25.8 ± 2.8
Highest language milestone attained												
All (24–41 months)	15.6 ± 4.0	15.7 ± 4.1	15.9 ± 4.1	15.2 ± 4.1	15.8 ± 3.9	15.3 ± 3.9	16.9 ± 3.6*	17.0 ± 3.6	17.3 ± 3.5	16.9 ± 3.6	16.8 ± 3.6	16.4 ± 3.8
24–29 months	13.5 ± 3.7	13.5 ± 3.8	13.8 ± 4.0	13.0 ± 3.8	13.6 ± 3.7	13.4 ± 3.5	14.9 ± 3.8*	15.0 ± 3.8	15.4 ± 3.8	14.9 ± 3.8	14.7 ± 3.7	14.0 ± 3.8
30–35 months	15.8 ± 3.8	16.0 ± 3.8	16.4 ± 3.7	15.2 ± 3.9	16.2 ± 3.8	15.4 ± 3.7	17.2 ± 3.3	17.3 ± 3.2	17.7 ± 3.1	17.2 ± 3.3	17.1 ± 3.2	16.6 ± 3.6
36–41 months	17.6 ± 3.3	17.8 ± 3.2	18.0 ± 3.3	17.8 ± 3.1	17.6 ± 3.2	17.3 ± 3.4	18.6 ± 2.6	18.7 ± 2.6	19.0 ± 2.3	18.5 ± 2.6	18.5 ± 2.8	18.3 ± 2.9

[a] Values are mean number of milestones ± SD. SDs adjusted for clustering.
[b] Sample size at baseline ranged from N = 1016 to 3545 in the full sample; N = 729 to 2557 in the treatment arms; N = 265 to 982 in the T24 arm; N = 182 to 583 in the T18 arm; N = 281 to 993 in the TNFP arm; and N = 287 to 988 in the control arm.
[c] Sample size at follow-up ranged from N = 1087 to 3530 in the full sample; N = 888 to 2959 in the treatment arms; N = 290 to 979 in the T24 arm; N = 302 to 990 in the T18 arm; N = 296 to 990 in the TNFP arm; and N = 164 to 571 in the control arm.
* Study arms differed, p < 0.05.

Table 65. Motor and language milestones attained: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Highest motor milestone attained				
All (24–41 months)	0.4 ± 0.3	0.5 ± 0.3	0.5 ± 0.4	0.2 ± 0.4
24–29 months	0.4 ± 0.3	0.5 ± 0.3*	0.7 ± 0.4	0.0 ± 0.4
30–35 months	0.4 ± 0.5	0.4 ± 0.5	0.7 ± 0.6	0.1 ± 0.6
36–41 months	0.4 ± 0.4	0.6 ± 0.5	0.2 ± 0.5	0.4 ± 0.5
Highest language milestone attained				
All (24–41 months)	0.3 ± 0.3	0.4 ± 0.4	0.6 ± 0.4	0.1 ± 0.4
24–29 months	0.9 ± 0.3*	1.0 ± 0.3*	1.5 ± 0.5*	0.5 ± 0.4
30–35 months	0.0 ± 0.5	0.0 ± 0.5	0.6 ± 0.6	-0.5 ± 0.6
36–41 months	0.1 ± 0.6	0.2 ± 0.7	-0.2 ± 0.6	0.1 ± 0.7

[a] Values are double-difference impact estimates ± SE in variable mean. All estimates control for child sex, child age, household size, household dependency ratio, mother's age, mother's education, household head education, and household head occupation at baseline and follow-up, and for *colline*-fixed effects. SEs adjusted for clustering.
[b] Sample size ranged from N = 2098 to 7027.
[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for all indicators.
* Impact estimate significantly different from 0, p < 0.05.

10.3 Child Hemoglobin Levels and Anemia Status

The mean Hb concentration among children 24–41 months of age at baseline was 10.9 g/dL, but this decreased to 10.6 g/dL at follow-up (**Table 66**). Correspondingly, the prevalence of anemia (Hb < 11 g/dL) increased from 47.9 percent to 56.6 percent. The most dramatic increase in anemia between baseline and follow-up was observed in children 36–41 months of age (from 45.7 percent to 58.5 percent, respectively). The prevalence of severe anemia (Hb < 7 g/dL) among all children 24–41 months increased from 1.8 percent to 2.2 percent from baseline to follow-up. At baseline, statistically significant differences between treatment groups for both Hb concentration (highest in the T24 group and lowest in the control group among all children and across all age groups) and the prevalence of anemia (lowest in the T24 group and highest in the control group among all children, and lowest in the TNFP group and highest in the control group among children 24–29 months of age) were observed.

No significant protective effects on child Hb levels and anemia were found among children 24–41 months of age (**Table 67**). The program had a 2.5 percentage point overall protective effect on severe anemia among children 36–41 months of age (significant in the T24 group); however, the prevalence of severe anemia was very low in all age groups.

Table 66. Child hemoglobin levels and anemia status[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	3475	2502	968	570	964	973	3482	2922	974	968	980	560
Child hemoglobin: adjusted value (g/dL)												
All (24–41 months)	10.9 ± 1.5*	11.0 ± 1.4	11.2 ± 1.3	10.7 ± 1.5	11.0 ± 1.5	10.7 ± 1.6	10.6 ± 1.6	10.6 ± 1.5	10.8 ± 1.4	10.5 ± 1.6	10.6 ± 1.5	10.3 ± 1.6
24–29 months	10.9 ± 1.5*	11.0 ± 1.4	11.1 ± 1.3	10.7 ± 1.5	11.0 ± 1.5	10.6 ± 1.6	10.7 ± 1.5	10.7 ± 1.5	10.9 ± 1.4	10.6 ± 1.6	10.6 ± 1.6	10.4 ± 1.6
30–35 months	10.9 ± 1.5*	11.0 ± 1.4	11.2 ± 1.3	10.8 ± 1.5	10.9 ± 1.5	10.7 ± 1.6	10.5 ± 1.6	10.6 ± 1.6	10.8 ± 1.4	10.4 ± 1.7	10.5 ± 1.6	10.3 ± 1.5
36–41 months	11.0 ± 1.5*	11.1 ± 1.5	11.2 ± 1.3	10.8 ± 1.6	11.1 ± 1.5	10.8 ± 1.5	10.6 ± 1.6	10.6 ± 1.5	10.7 ± 1.5	10.5 ± 1.6	10.6 ± 1.5	10.3 ± 1.6
Child anemia[d]												
All (24–41 months)	47.9*	45.6	43.2	51.4	44.5	53.9	56.6	55.4	51.8	57.7	56.6	62.7
24–29 months	49.1*	46.0	44.1	53.3	43.5	57.5	53.6	52.6	49.9	54.2	53.8	59.6
30–35 months	48.9	46.9	39.5	52.5	50.2	54.0	57.5	56.3	51.7	58.5	58.7	62.9
36–41 months	45.7	44.0	45.1	48.0	41.0	49.8	58.5	57.3	54.0	60.2	57.6	65.1
Child severe anemia[e]												
All (24–41 months)	1.8	1.5	0.9	2.5	1.5	2.5	2.2*	1.9	1.2	3.0	1.5	3.6
24–29 months	1.9	1.5	1.1	2.8	1.2	2.8	2.4	1.9	1.5	2.9	1.5	4.8
30–35 months	2.1	1.4	0.8	1.7	1.8	3.9	1.9	1.9	1.0	3.3	1.3	2.0
36–41 months	1.3	1.5	0.9	2.9	1.4	0.9	2.3	1.9	1.2	2.8	1.7	4.2

[a] Values are % or mean ± SD. SDs adjusted for clustering.

[b] Sample size at baseline ranged from N = 1000 to 3475 in the full sample; N = 715 to 2502 in the treatment arms; N = 261 to 968 in the T24 arm; N = 175 to 570 in the T18 arm; N = 273 to 964 in the TNFP arm; and N = 285 to 973 in the control arm.

[c] Sample size at follow-up ranged from N = 1104 to 3482 in the full sample; N = 902 to 2922 in the treatment arms; N = 296 to 974 in the T24 arm; N = 306 to 968 in the T18 arm; N = 300 to 980 in the TNFP arm; and N = 166 to 560 in the control arm.

[d] Cutoff was 11 g/dL.

[e] Cutoff was 7 g/dL.

* Study arms differed, $p < 0.05$.

Table 67. Child hemoglobin levels and anemia status: post-program impact[a]

	Post-program impact[b,c]			
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control
Child hemoglobin: adjusted value (g/dL)				
All (24–41 months)	0.0 ± 0.1	-0.1 ± 0.2	0.1 ± 0.2	-0.1 ± 0.1
24–29 months	-0.1 ± 0.2	-0.1 ± 0.2	0.1 ± 0.2	-0.1 ± 0.2
30–35 months	0.0 ± 0.2	0.0 ± 0.2	0.1 ± 0.3	0.0 ± 0.2
36–41 months	0.0 ± 0.2	0.0 ± 0.2	0.1 ± 0.2	-0.1 ± 0.2
Child anemia				
All (24–41 months)	0.6 ± 3.7	0.3 ± 5.1	-2.9 ± 4.1	3.5 ± 3.8
24–29 months	3.0 ± 4.4	2.8 ± 5.0	-2.0 ± 5.3	6.7 ± 5.6
30–35 months	-2.5 ± 4.8	0.7 ± 7.8	-6.8 ± 5.8	-2.3 ± 5.1
36–41 months	-1.0 ± 4.9	-5.1 ± 6.5	-1.7 ± 6.0	3.5 ± 5.0
Child severe anemia				
All (24–41 months)	-0.8 ± 0.7	-0.8±0.8	-0.6 ± 1.1	-1.0 ± 0.9
24–29 months	-2.1 ± 1.7	-1.9±1.9	-2.3 ± 2.1	-2.1 ± 1.8
30–35 months	1.2 ± 1.7	1.3±1.5	2.1 ± 2.3	0.4 ± 1.9
36–41 months	-2.5 ± 1.1*	-2.5±1.1*	-2.8 ± 1.8	-2.3 ± 1.6

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points. All estimates control for altitude, child sex, child age, household size, household dependency ratio, mother's age, mother's education, household head education, and household head occupation at baseline and follow-up, and for *colline*-fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 2096 to 6910.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for mean Hb level and a negative effect for anemia and severe anemia.

* Impact estimate significantly different from 0, $p < 0.05$.

10.4 Child Anthropometry

At baseline, the average HAZ was extremely low (-2.3 SD) and the corresponding prevalence of stunting alarmingly high (overall 65.2 percent) (**Table 68**). A large deterioration of the situation was observed in the control group (HAZ decreased from -2.4 to -2.6 and the stunting prevalence increased from 68.2 percent to 74.8 percent). The groups were balanced at baseline (except for stunting in children 30 to 35 months of age).

The average WHZ for the full sample at baseline was -0.1 SD, and the corresponding prevalence of wasting was 3.0 percent (**Table 68**). At baseline and follow-up, WHZ and the prevalence of wasting were similar within the full sample and each age group. Wasting was noticeably higher among children 24–29 months at baseline and follow-up. There were no significant differences in wasting or WHZ at baseline.

Tubaramure had a large post-program protective effect on stunting of 5.5 percentage points (all arms combined). A significant effect was also seen in the T24 group (7.1 percentage points), but did not reach significance in the T18 or TNFP groups (**Table 69**). The central research question of the Burundi study was to determine the optimal program duration to improve linear growth. Even though the impact was not statistically significant between treatment arms, our findings suggest that program exposure from pregnancy to 24 months was necessary for impact on linear growth in the beneficiary population in Burundi.

Analyses by age group showed that the post-program impact was limited to children 24–29 months of age, in which the *Tubaramure* program had an overall estimated protective effect of 16.3 percentage points. This significant effect was found in all treatment groups (16.6, 10.2, and 20.4 percentage points in the T24, T18, and TNFP arms, respectively). Similarly, the intervention had an overall significant impact on HAZ in this age group (of 0.3 SD), which was also found in the T24 (0.3 SD) and TNFP (0.4 SD) groups.

Tubaramure did not have a post-program protective effect on wasting or WHZ.

Table 68. Child stunting and wasting[a]

	Baseline[b]						Follow-up[c]					
		Study arm						Study arm				
	Full sample	All treated	T24	T18	TNFP	Control	Full sample	All treated	T24	T18	TNFP	Control
N	3453	2495	959	565	971	958	3500	2940	974	982	984	560
Child HAZ												
All (24–41 months)	-2.3 ± 1.2	-2.3 ± 1.2	-2.3 ± 1.1	-2.3 ± 1.2	-2.3 ± 1.1	-2.4 ± 1.2	-2.4 ± 1.1*	-2.3 ± 1.1	-2.3 ± 1.1	-2.3 ± 1.1	-2.4 ± 1.1	-2.6 ± 1.0
24–29 months	-2.4 ± 1.2	-2.4 ± 1.2	-2.3 ± 1.1	-2.2 ± 1.3	-2.5 ± 1.1	-2.5 ± 1.2	-2.3 ± 1.1*	-2.2 ± 1.1	-2.2 ± 1.2	-2.3 ± 1.1	-2.3 ± 1.1	-2.6 ± 1.0
30–35 months	-2.4 ± 1.1	-2.4 ± 1.1	-2.3 ± 1.1	-2.4±1.2	-2.4±1.1	-2.6±1.1	-2.4 ± 1.1	-2.4 ± 1.1	-2.4 ± 1.0	-2.3 ± 1.1	-2.4 ± 1.1	-2.6 ± 1.0
36–41 months	-2.2 ± 1.2	-2.2 ± 1.1	-2.2 ± 1.1	-2.1 ± 1.2	-2.2 ± 1.1	-2.3 ± 1.2	-2.4 ± 1.0	-2.4 ± 1.0	-2.3 ± 1.1	-2.4 ± 1.1	-2.4 ± 1.0	-2.6 ± 1.0
Child stunted												
All (24–41 months)	65.2	64.1	64.7	62.3	64.6	68.2	65.9*	64.3	63.0	64.1	65.7	74.8
24–29 months	67.3	67.5	67.4	62.4	70.7	66.9	64.4*	61.9	61.7	63.3	60.6	79.4
30–35 months	67.8*	65.0	64.6	68.0	63.4	74.9	67.3	66.3	66.2	63.4	69.2	71.7
36–41 months	60.8	59.6	61.5	56.3	59.5	63.8	66.2	64.8	61.6	65.3	67.3	74.2
Child WHZ												
All (24–41 months)	-0.1 ± 1.0	-0.1 ± 1.0	-0.2 ± 1.0	-0.2 ± 1.1	-0.1 ± 1.0	0.0 ± 1.0	0.0 ± 1.0	0.0 ± 1.0	-0.1 ± 0.9	0.0 ± 1.1	0.0 ± 1.0	0.0 ± 0.9
24–29 months	-0.2 ± 1.0	-0.3 ± 1.0	-0.2 ± 0.9	-0.3 ± 1.1	-0.3 ± 1.0	-0.2 ± 1.1	-0.1 ± 1.0	-0.1 ± 1.0	-0.1 ± 0.9	-0.1 ± 1.1	-0.1 ± 1.0	-0.1 ± 1.0
30–35 months	0.0 ± 1.0	0.0 ± 1.0	0.0 ± 1.1	0.0 ± 1.0	-0.1 ± 0.9	0.0 ± 1.1	0.0 ± 1.0	0.0 ± 1.0	-0.1 ± 1.0	0.0 ± 1.1	0.0 ± 1.0	0.0 ± 0.9
36–41 months	-0.1 ± 1.0	-0.1 ± 1.0	-0.1 ± 1.0	-0.1 ± 1.0	0.0 ± 1.0	0.1 ± 1.0	0.0 ± 0.9	0.0 ± 1.0	-0.1 ± 0.9	0.1 ± 1.0	-0.1 ± 1.0	0.1 ± 0.9
Child wasted												
All (24–41 months)	3.0	3.1	2.4	3.9	3.3	2.9	2.6	2.6	2.3	2.9	2.7	2.3
24–29 months	3.8	3.7	2.7	5.6	3.8	4.0	3.3	3.0	2.1	3.8	3.3	4.8
30–35 months	2.2	1.8	1.6	1.7	2.2	3.1	2.1	2.3	2.4	2.9	1.7	1.0
36–41 months	2.9	3.4	2.8	4.0	3.7	1.5	2.3	2.5	2.3	2.0	3.1	1.6

[a] Values are % or mean ± SD. SDs adjusted for clustering.

[b] Sample size at baseline ranged from N = 964 to 3469 in the full sample; N = 694 to 2500 in the treatment arms; N = 253 to 959 in the T24 arm; N = 171 to 568 in the T18 arm; N = 267 to 974 in the TNFP arm; and N = 270 to 969 in the control arm.

[c] Sample size at follow-up ranged from N = 1093 to 3508 in the full sample; N = 891 to 2944 in the treatment arms; N = 289 to 975 in the T24 arm; N = 302 to 982 in the T18 arm; N = 300 to 987 in the TNFP arm; and N = 160 to 564 in the control arm.

* Study arms differed, $p < 0.05$.

Table 69. Child stunting and wasting: post-program impact[a]

	Post-program impact[b,c]					
	T vs. control	T24 vs. control	T18 vs. control	TNFP vs. control	T24 vs. T18	T24 vs. TNFP
Child HAZ						
All (24–41 months)	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.0 ± 0.1	0.0 ± 0.1
24–29 months	0.3 ± 0.1*	0.3 ± 0.1*	0.2 ± 0.1	0.4 ± 0.1*	-0.1 ± 0.1	0.0 ± 0.1
30–35 months	0.0 ± 0.1	-0.1 ± 0.1	0.1 ± 0.2	-0.1 ± 0.2	0.2 ± 0.1	0.1 ± 0.2
36–41 months	0.0 ± 0.1	0.1 ± 0.1	-0.1 ± 0.1	-0.1 ± 0.2	-0.2 ± 0.1	-0.2 ± 0.2
Child stunted						
All (24–41 months)	-5.5 ± 2.7*	-7.1 ± 3.2*	-5.2 ± 3.2	-4.0 ± 3.3	2.0 ± 3.1	3.1 ± 3.2
24–29 months	-16.3 ± 3.6*	-16.6 ± 4.5*	-10.2 ± 4.5*	-20.4 ± 5.0*	6.4 ± 5.0	-3.8 ± 5.4
30–35 months	3.5 ± 4.5	3.4 ± 6.1	-1.3 ± 5.5	7.5 ± 6.0	-4.7 ± 6.4	4.1 ± 6.7
36–41 months	-3.3 ± 5.1	-9.0 ± 6.0	-0.2 ± 5.7	0.0 ± 7.2	8.7 ± 5.8	9.0 ± 7.3
Child WHZ						
All (24–41 months)	0.1 ± 0.1	0.0 ± 0.1	0.1 ± 0.1	0.0 ± 0.1	–	–
24–29 months	0.1 ± 0.1	0.0 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	–	–
30–35 months	0.0 ± 0.1	-0.1 ± 0.1	0.0 ± 0.2	0.1 ± 0.1	–	–
36–41 months	0.0 ± 0.1	0.0 ± 0.1	0.3 ± 0.1	-0.1 ± 0.1	–	–
Child wasted						
All (24–41 months)	-0.1 ± 1.0	0.3 ± 1.2	-0.6 ± 1.3	0.0 ± 1.2	–	–
24–29 months	-1.5 ± 1.5	-0.7 ± 1.7	-3.2 ± 2.4	-1.0 ± 1.8	–	–
30–35 months	2.7 ± 2.1	3.0 ± 2.2	3.7 ± 2.6	1.8 ± 2.2	–	–
36–41 months	-1.0 ± 1.5	-0.8 ± 2.0	-2.0 ± 1.8	-0.6 ± 2.1	–	–

[a] Values are double-difference impact estimates ± SE in variable mean or percentage points. All estimates control for child sex, child age, household size, household dependency ratio, mother's age, mother's education, household head education, and household head occupation at baseline and follow-up, and for *colline*-fixed effects. SEs adjusted for clustering.

[b] Sample size ranged from N = 2037 to 6890.

[c] One-sided tests were conducted. The *a priori* hypothesis was a positive effect for mean height-for-age difference and HAZ, and a negative effect for stunting.

* Impact estimate significantly different from 0, $p < 0.05$.

11. Summary of Findings along the Program Impact Pathways

In this section of the report, we first summarize (past) program participation and then discuss the key post-program impacts along the primary program impact pathways that correspond to the three program components: the food component, the health component, and the BCC component. Where relevant, we also provide a summary of findings from the 2012 follow-up survey. We conclude with a section on the post-program impacts on maternal and child Hb concentrations and anemia, and on child linear growth and development.

11.1 *Tubaramure* Program Participation

At the time of the 2014 follow-up, all program activities, including the food distribution, had ended. A large percentage (67 percent to 77 percent) of respondents reported being previous program beneficiaries. This is around 8 percentage points lower than what was reported in the 2012 survey, when the program was ongoing (Leroy et al. 2014). The lower reported participation in the 2014 survey might be due to migration, which would have reduced the percentage of respondents in the treatment *collines* who were exposed to the program. Recall bias (i.e., former beneficiaries have forgotten the name of the program) might be another explanation.

11.2 The *Tubaramure* Food Component

The food component of the *Tubaramure* program was expected to increase household availability of micronutrient-rich food and, in turn, consumption of such foods, and to improve diet diversity. All *Tubaramure* beneficiaries were entitled to a monthly household food ration and an individual food ration meant to be consumed by the woman while she was pregnant or when her child was under 6 months of age and by the child when he or she was 6–23 months of age.

At the first follow-up survey, conducted in 2012 when *Tubaramure* was still ongoing, participation in food distribution was high (Leroy et al. 2014). The program had a large positive effect on household access to food and substantially reduced the percentage of severely food insecure households. The impact on the low level of household and mothers' dietary diversity, however, was minimal. The program improved complementary feeding practices,[55] with the largest impact seen in the percentage of children consuming iron-rich foods. This effect appeared to be directly related to the increased intake of CSB. A more modest positive impact was found for the percentage of children receiving the minimum recommended number of meals, the percentage of children receiving the minimum dietary diversity, and the percentage of children consuming a minimally acceptable diet. The remainder of this section summarizes the 2014 post-program impacts.

Tubaramure had a positive lasting effect on household access to food as measured by the HFIAS (Coates et al. 2007). The program increased the percentage of food secure households by 4 to 6 percentage points. *Tubaramure* also increased household dietary diversity by 0.3 food groups and reduced the percentage of households that consumed fewer than four food groups (out of a possible 12) by around 12 percentage points. The program had a similar positive post-program impact of 0.4 food groups (out of a possible nine groups) on diversity of mothers' diets. The largest impact for a single food group was found for non-vitamin A-rich fruits and vegetables.

There are different possible pathways by which *Tubaramure* could have had these longer-term effects (i.e., effects after the program ended). One possibility is that households used some of the freed-up

---

[55] Complementary feeding practices were assessed in children 6–23 months of age. The 2014 survey focused on children 24–41 months of age, so complementary feeding practices were not assessed.

resources to invest in productive assets, which consequently led to higher household income and expenditure, and ultimately to improvements in food security and dietary diversity. In addition, the program, through its BCC component, could have changed beneficiaries' understanding of the importance of healthy diets, which in turn could have led to an increase in the number of food groups consumed. Assessing the (relative) strength of these pathways,[56] however, is beyond the scope of this report.

11.3 The *Tubaramure* Health Component

Tubaramure's health service strengthening component aimed to improve the quality of preventive and curative care provided at public health centers. The intervention was implemented in all health centers regardless of whether they were located in a treatment or control *colline*; the lack of a control group thus prevents attributing any changes in health center characteristics directly to *Tubaramure* (see Section 2.4.7). Differences between baseline and follow-up suggest that several aspects of health service delivery may have improved.

In 2012, health centers were better equipped and better staffed to care for expecting mothers and young children, but shortages of key immunizations and medications were common.[57] The 2014 findings are summarized below.

At follow-up, all but one of the health centers staffed A2-level nurses (the highest qualified staff members mandated for health centers), which was more than the five centers that staffed A2-level nurses at baseline. Three centers were also staffing an A2-level laboratory technician at follow-up.

Preventive services for children under 5 were nearly universally available at baseline, and with the exception of vitamin A supplementation and deworming, this was also the case at follow-up. Key service components for sick children, children with diarrhea, and pregnant women improved over time. Specifically, health education, the evaluation of preventive care status (i.e., chart weight, immunizations, and vitamin A supplementation every 6 months), and services for the treatment of diarrhea for children (i.e., oral serum, onsite intravenous solution, and zinc supplementation) improved dramatically. For women seeking prenatal care, there were improvements in the availability of services such that, with the exception of niclosamide for *taenia* deworming, all surveyed components of services for pregnant and lactating women are now universally available in all 11 health centers.

HIV testing for pregnant women became available at every health center at follow-up; however, urine and anemia tests for pregnant women were available at fewer than half of the health centers. Blood and stool tests for children under 5 were available at all but one health center, but urine tests were offered in only one health center. Blood tests for malaria in children were universally available.

All health centers generally stocked key vaccines, micronutrients, and supplements, but stock ruptures were common. At least one health center reported ruptures in the supply chain of every key vaccine except BCG and measles, and nearly all health centers that stocked iron syrup reported shortages. For the treatment of infections, antibiotics were generally available. Analgesics, oral contraceptives, antimalarial medications, and anthelmintics (with the exception of praziquantel and niclosamide) were also readily available.

---

[56] Another possible pathway is community gardens, which *Tubaramure* promoted through the BCC component. This program component, however, was introduced too late to assess its coverage and impact.

[57] Note that *Tubaramure*, under FFP guidelines, was not able to purchase immunizations or medications.

11.4 The *Tubaramure* Behavior Change Communication Component

Tubaramure's BCC strategy was designed to encourage the adoption of best practices in health, hygiene, and nutrition. It was designed to be implemented by program staff and locally hired THPs. The THPs trained leader mothers, who were program beneficiaries selected by their fellow beneficiary mothers. Leader mothers in turn trained the beneficiary mothers in bi-weekly care groups on the best practices and topics most recently learned from the THP.

The THPs trained leader mothers who were program beneficiaries selected by their fellow beneficiary mothers to teach them. Leader mothers in turn trained the beneficiary mothers in bi-weekly care groups on the topics most recently learned from the THP.

The 2012 follow-up survey, conducted when the program was still ongoing, found that care group participation was low: mothers were exposed to only about half of the intended BCC sessions. The process evaluation (conducted from late 2011 to early 2012) revealed that *Tubaramure*'s BCC strategy suffered substantial delays in the design and rollout of the complementary feeding module. As a result, only a fraction of mothers had been exposed to this module at the time of the 2012 follow-up survey. The process evaluation further showed that the leader mothers had limited technical expertise and teaching skill (Olney et al. 2013).

The impact on mothers' knowledge in 2012 was mixed; mothers' knowledge did not improve in many of the areas in which it was expected to improve (Leroy et al. 2014). Mothers' knowledge did improve in the area of optimal child feeding practices, but *Tubaramure* had only a modest positive impact on a limited number of household hygiene and sanitation practices. No impact was found on reported attendance at growth monitoring services or utilization of any of the other types of preventive health care services. The program did protect children from the overall increase in child morbidity observed from baseline to the 2012 follow-up in the control group. The only effect on curative care seeking behaviors was for the percentage of children receiving medication to fight a fever. We concluded from the 2012 survey that better delivery and utilization of the BCC strategy might have resulted in a larger impact on maternal knowledge, health and nutrition practices, and nutritional status. The remainder of this section summarizes the post-program impacts as documented in the 2014 follow-up survey.

11.4.1 Health and Nutrition Knowledge

The program's post-program effect on mothers' very limited knowledge of danger signs for childhood illnesses was limited to a small increase in the percentage of mothers mentioning rapid breathing. *Tubaramure* had a lasting effect on mother's knowledge of breastfeeding, providing liquids, and appropriate feeding for sick children; no post-program effect was found for knowledge regarding feeding a child recovering from illness.

Tubaramure, a program aimed at improving child nutritional status, had no clear post-program impact on mothers' limited knowledge of the reasons for child malnutrition. The program greatly increased awareness of vitamin A-rich foods and had a modest post-program effect on mothers who mentioned impaired vision as a consequence of vitamin A deficiency. The program did not change the percentage of mothers who knew the consequences of iron deficiency, however, and the post-program impact on awareness regarding iron-rich foods was limited to CSB.

The *Tubaramure* program had positive post-program effects on knowledge regarding the age of introduction of liquids and solid foods other than breast milk: it increased the percentage of mothers who knew that liquids (other than breast milk) and foods should not be introduced before 6 months of age. The program did not, however, reduce the percentage of mothers who wrongly believed that complementary foods should be introduced after 6 months of age. *Tubaramure's* post-program impact on other feeding-related knowledge (e.g., consistency of complementary foods, feeding frequency, eating without help) was limited or absent.

The program had large positive post-program effects on many of the maternal hygiene knowledge outcomes. It dramatically increased the percentage of mothers who could mention key handwashing times. A more modest post-program effect was observed for maternal awareness of appropriate worm-protection methods. No significant effect was observed for knowledge of appropriate methods for purifying drinking water.

In summary, *Tubaramure*'s post-program impact on maternal health and nutrition knowledge was mixed, with clear improvements in handwashing and some knowledge of nutrient-rich foods (e.g., vitamin A-rich foods), more modest improvements in knowledge related to specific optimal practices (e.g., feeding frequency), and no detectable effect in areas in which key health and nutrition knowledge would have ideally improved (e.g., danger signs of child illness, reasons for undernutrition).

11.4.2 Hygiene and Sanitation Practices

Tubaramure had a modest positive post-program impact on household hygiene and sanitation practices. The percentage of households that treated their water increased only slightly, but a considerable post-program effect was found on the percentage of households that composted their trash rather than discarding it in a public space. The importance of composting was emphasized on a *Tubaramure* household poster provided to all beneficiaries. Bed net ownership and use increased dramatically between baseline and follow-up across all study arms; no post-program impact on bed net ownership or use was found. Even though most households owned soap, it was rarely used at key handwashing times; the post-program effect of *Tubaramure* was limited to a small increase in the percentage of households reporting the use of soap after defecating and after cleaning a child that had defecated. *Tubaramure* was not found to improve the cleanliness of the hands, hair, clothes, and face of mothers and children (based on spot checks), nor did it improve the cleanliness of household interiors. The percentage of households with clean exteriors slightly improved.

11.4.3 Impact on Preventive and Curative Care Seeking and Child Health

Tubaramure did not have a post-program impact on the utilization of growth monitoring services, on the percentage of children having received vitamin A supplementation in the 6 months preceding the survey, or on the percentage of children being fully vaccinated.

However, *Tubaramure*, even after the program ended, protected children from illness: the increase in reported symptoms from baseline to follow-up in the control group was not observed in the *Tubaramure* groups; many of the morbidity symptoms in the treatment groups actually declined over time. Significant post-program effects were found for general morbidity symptoms and for key problems, such as fever, convulsions, difficulty drinking, loss of appetite, cough, difficulty breathing, and watery diarrhea.

The program did not have any significant positive post-program impacts on curative care seeking behaviors, which were found to be inadequate: at follow-up, only around 40 percent of children with diarrhea had received ORS, and approximately 30 percent of mothers reported reducing the child's liquid intake during diarrhea. The percentage not seeking care for illness requiring immediate medical attention dropped from around one-third to around 22 percent over time, but this decline was not larger in the *Tubaramure* study arms.

11.5 Impact on Maternal Nutrition and Child Nutrition and Development

The three *Tubaramure* components discussed above (i.e., food, health, and BCC) were expected to have impacts on maternal nutrition and on child nutrition and development. The first follow-up survey documented a steep increase in the prevalence of anemia in mothers from 2010 to 2012. The program had a significant protective effect on maternal anemia of 4.2 to 7.5 percentage points. A similar result was found in children 6–23 months of age, with an overall increase from 2010 to 2012 in the prevalence of anemia (to around 73 percent) and with the program having a protective effect on children's Hb of 0.24 to 0.56 g/dL. A general decline between surveys in the highest attained language and motor milestones among children 4–23 months of age was also observed. *Tubaramure* had a protective effect on the highest attained motor milestone and on the highest attained language milestone in some of the intervention groups. The 2014 post-program impacts on these outcomes are summarized in the following sections.

11.5.1 Hemoglobin Levels and Anemia

There was an overall decrease in mothers' mean Hb levels between the 2010 and 2014 surveys, corresponding to a 12.2 percentage point increase in anemia. *Tubaramure* did not have a post-program protective effect on Hb levels or anemia status.

A similar decline in mean Hb levels (by 0.3 g/dL) and an increase in the prevalence of anemia (by 8.7 percentage points) was observed in children. No significant protective effects from the intervention on child Hb levels and anemia (Hb < 11 g/dL) were found. However, there was a large post-program protective effect of 2.5 percentage points on the reduction of severe anemia (Hb < 7 g/dL) in both the combined treatment and T24 groups, among children 36–41 months of age.[58]

11.5.2 Impact on Child Development and Child Linear growth

Both the mean highest attained motor milestone and the mean highest attained language milestone increased between surveys. The post-program impact on the highest motor milestone achieved was limited to a 0.5 milestone effect in children 24–29 months of age in the T24 group. *Tubaramure* was found to have an overall positive post-program effect on the highest language milestones achieved among children 24–29 months of age.

The second follow-up survey found a steep increase in the already high prevalence of stunting in the control group (68.2 percent to 74.8 percent). *Tubaramure* had a large post-program protective effect on stunting of 5.5 percentage points (all treatment groups combined). A significant effect was also seen in the T24 group (7.1 percentage points), but did not reach significance in the T18 or TNFP groups (**Table 69**). Even though the difference in impact was not statistically significant between treatment arms, our findings suggest that program exposure from pregnancy to 24 months was necessary for impact on linear growth in the beneficiary population in Burundi. This finding answers the central research question of the Burundi study, which was to determine the optimal program duration to improve linear growth.

It seems plausible that *Tubaramure*'s effect on improving infant and young child feeding practices (i.e., minimum recommended number of meals, the percentage of children receiving the minimum dietary diversity, the percentage of children consuming a minimally acceptable diet) and on reducing child morbidity (Leroy et al. 2014) contributed to this linear growth impact. Identifying and quantifying the exact impact pathways, however, is beyond the scope of this report.

[58] Note that the prevalence of severe anemia was low: at follow-up, it was 1.9 percent and 4.2 percent in the treatment and control arms, respectively.

Analyses by age group showed that the largest linear growth impact was found in children 24–29 months of age, in which the *Tubaramure* program had an overall estimated protective effect of 16.3 percentage points. No significant post-program impact was found in the other age groups. The prevalence of wasting in children at the second follow-up was low and *Tubaramure* did not have a post-program effect on this outcome.

There are different possible explanations why *Tubaramure*'s effect on the child development and child linear growth was (largely) limited to children 24–29 months of age. A first obvious possibility is that program participation was higher in this age group than in older age groups. Our data show, however, that this was not the case. Second, the quality of program implementation (which was not measured in the impact evaluation surveys) might have improved over time as program staff gained experience. If this was the case, then the youngest children in the 2014 follow-up survey would have benefited from higher-quality program services more than older children in the sample. A third possibility is that the program's impact could not be sustained when the program benefits were no longer received. Future analyses of the data will focus on better understanding this age effect.

12. References

Ballard, T.,Coates, J., Swindale, A., and Deitchler, M. 2011. *Household Hunger Scale: Indicator Definition and Measurement Guide.* Washington, DC: FANTA-2 Bridge, FHI 360.

Central Intelligence Agency. 2016. *CIA World Factbook*. Washington, DC.

Coates, J., Swindale, A., and Bilinsky, P. 2007. *Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide (v.3).* Washington, DC: FHI 360/FANTA.

Cogill, B. 2003. *Anthropometric Indicators Measurement Guide*. Washington, DC: Food and Nutrition Technical Assistance (FANTA) Project, FHI 360.

Deitchler, M., Ballard, T., Swindale, A., and Coates, J. (2010). *Validation of a Measure of Household Hunger for Cross-Cultural Use*. Washington, DC: Food and Nutrition Technical Assistance II Project (FANTA-2), FHI 360.

EICV. 1998. *Rwanda Enquête Intégral sur les Conditions de Vie des Ménages (EICV)*.

Griffiths, R. 1970. *A Comprehensive System of Mental Measurement for the First Eight Years of Life*. London: Child Development Research Centre.

International Monetary Fund. 2012. *Burundi: Poverty Reduction Strategy Paper II. IMF Country Report Number 12/224.* Washington, DC: Internatinoal Monetary Fund.

ISTEEBU and INSP. 2011. *Enquête Démographique et de Santé Burundi 2010 - Rapport Préliminaire*. Bujumbura, Burundi: ISTEEBU, MSPLS, and ICF International.

International Fund for Agricultural Development. 2012. *Enabling Poor Rural People to Overcome Poverty in Burundi*. Rome: International Fund for Agricultural Development.

Nestel, P. 2002. *Adjusting Hemoglobin Values in Program Survey*. Washington, DC: International Nutritional Anemia Consultative Group.

Kennedy, G., Ballard, T., and Dop, M. 2011. *Guidelines for Measuring Household and Individual Dietary Diversity.* Rome: Food and Agriculture Organization.

Leroy, J. L., Heckert, J., Cunningham, K., and Olney, D.K. 2014. *Strengthening and Evaluating the Preventing Malnutrition in Children under 2 Years of Age Approach–Burundi Follow-Up Report: Children 0–23 Months*. Washington, DC: FHI 360/FANTA.

McCarthy, D. 1972. *Manual for the McCarthy Scales of Children's Abilities*. New York: Psychological Corporation.

Ministère de la Santé Publique et de la Lutte contre le SIDA. 2012. *La Liste Nationales des Médicaments Essentiels*. 7th edition. Bujumbura, Burundi.

Ministère de la Santé Publique et de la Lutte contre le Sida. 2012. *Normes Sanitaires pour la Mise en Oeuvre Dupnds 2011-2015 au Burundi*. Bujumbura, Burundi.

Olney, D.K., et al. 2009. "Development of Nutritionally At-Risk Young Children Is Predicted by Malaria, Anemia, and Stunting in Pemba, Zanzibar." *Journal of Nutrition.* 139(4): 763–772.

Olney, D.K., et al. 2013. *A Process Evaluation of the Tubaramure Program for Preventing Malnutrition in Children under 2 Approach (PM2A) in Burundi*. Washington, DC: FHI 360/FANTA.

Parker, M., et al. 2012. *Strengthening and Evaluating the Preventing Malnutrition in Children under 2 Approach (PM2A) in Burundi: Baseline Report*. Washington, DC: FHI 360/FANTA.

Rao, J. and Scott, A. 1984. "On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated from Survey Data." *Annals of Statistics.* 12(1): 6–60.

StataCorp. 2009. *Stata Statistical Software: Release 11.* College Station, TX: StataCorp LP.

Stevens, G.A., et al. 2013. "Global, Regional, and National Trends in Haemoglobin Concentration and Prevalence of Total and Severe Anaemia in Children and Pregnant and Non-Pregnant Women for 1995–2011: A Systematic Analysis of Population-Representative Data." *The Lancet Global Health*. 1(1)e: 16–e25.

Stoltzfus, R.J., et al. 2001. "Effects of Iron Supplementation and Anthelmintic Treatment in Zanzibar: Double Blind, Placebo Controlled Study." *British Medical Journal*. 323: 1389–1393.

Swindale, A. and Bilinsky, P. 2006. *Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide Version 2*. Washington, DC: FHI 360/FANTA.

UN Stats. 2009. *Recensement General de la Population et de l'Habitat 2008: Resultats Preliminaires Commentes*. Bujumbura, Burundi.

Unite Nations Development Programme. 2015. *Human Development Report 2015. Work for Human Development.* New York: United Nations Development Programme.

United Nations Department of Economic and Social Affairs Statistics Division. 2005. *Household Sample Surveys in Developing and Transition Countries (No. 96).* New York: United Nations.

WHO. 2008. *Indicators for Assessing Infant and Young Child Feeding Practices: Part 1 Definitions*. Geneva: WHO.

WHO. 2009. *Global Prevalence of Vitamin A Deficiency in Populations at Risk 1995-2005: WHO Global Database on Vitamin A Deficiency*. Geneva: WHO.

WHO. 2011a. *Hemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity.* Geneva: Vitamin and Mineral Nutrition Information System, WHO.

WHO. 2011b. *Iron Deficiency Anaemia Assessment, Prevention, and Control: A Guide for Programme Managers.* Geneva: WHO.

WHO Expert Committee on Physical Status. 1995. *Physical Status: The Use and Interpretation of Anthropometry.* Geneva: WHO.

WHO Multicentre Growth Reference Study Group. 2006. "WHO Child Growth Standards." *Acta Paediatr Suppl*. 95(S450): 1–104.

WHO Multicentre Growth Reference Study Group. 2006. *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development.* Geneva: WHO.

Wooldridge, J.M. 2003. *Introductory Econometrics: A Modern Approach (Vol. 2nd).* Australia and Cincinnati, Ohio: South-Western College Pub.

World Bank. 2016. *World Bank Development Indicators 2016*. Washington, DC: World Bank.

Appendix 1. Monthly Ration Size for the *Tubaramure* Beneficiary Population

	CSB (kg)	Vegetable oil (g)	Energy/ month (kcal)	Energy/day[a] (kcal)	Energy/day/ capita[b] (kcal)
Target group[c]					
Pregnant/lactating women	6	600	27,846	915	158[c]
Children under 2	3	300	13,923	458	79[c]
Family ration	12	1,200	55,692	1,831	316
Total ration					
Pregnant/lactating women	18	1,800	83,538	2,746	474
Children under 2	15	1,500	69,615	2,289	395

[a] Energy per day is calculated using 30.42 days/month.

[b] Energy per capita is calculated based on the assumption of average household size of 5.8 members.

[c] Note that the individual ration is not meant to be shared, but this may be difficult to achieve, as the individual and family rations include the same foods.

Appendix 2. List of *Collines* according to Study Treatment Arm

Province	Commune	*Colline*	Study arm[a]
Cankuzo	Cankuzo	Kabezera	T18
Cankuzo	Cankuzo	Murehe	TNFP
Cankuzo	Cankuzo	Muterero	T24
Cankuzo	Cankuzo	Nyarutiti	T24
Cankuzo	Cendajuru	Gashirwe	Control
Cankuzo	Cendajuru	Gitaramuka	TNFP
Cankuzo	Cendajuru	Kabageni	TNFP
Cankuzo	Cendajuru	Nyamugari	T18
Cankuzo	Gisagara	Gerero	T18
Cankuzo	Gisagara	Gisagara	T24
Cankuzo	Gisagara	Gitwenge	TNFP
Cankuzo	Gisagara	Kibogoye	T24
Cankuzo	Gisagara	Murago	Control
Cankuzo	Kigamba	Rujungu	T18
Cankuzo	Kigamba	Rwamvura	TNFP
Cankuzo	Kigamba	Shinge	T24
Cankuzo	Mishiha	Buyongwe 1	Control
Cankuzo	Mishiha	Mwiruzi	Control
Cankuzo	Mishiha	Rugerero	Control
Cankuzo	Mishiha	Rutsindu	T18
Ruyigi	Butaganzwa	Gikwiye	TNFP
Ruyigi	Butaganzwa	Kanyinya	TNFP
Ruyigi	Butaganzwa	Kirangara	Control
Ruyigi	Butaganzwa	Kiyabu	T24
Ruyigi	Butaganzwa	Masazi	T24
Ruyigi	Butaganzwa	Muriza	T18
Ruyigi	Butaganzwa	Nyagashubi	T24
Ruyigi	Butaganzwa	Taba	T18
Ruyigi	Butezi	Bwagiriza	T24
Ruyigi	Butezi	Muyange	TNFP
Ruyigi	Butezi	Nombe	T24
Ruyigi	Butezi	Rutegama	Control
Ruyigi	Bweru	Gatwaro	TNFP
Ruyigi	Bweru	Kirambi	T18
Ruyigi	Bweru	Mubavu	T24
Ruyigi	Bweru	Nkanda	T24
Ruyigi	Bweru	Nyarunazi	T24
Ruyigi	Gisuru	Bunyambo	T24
Ruyigi	Gisuru	Kabingo	T18
Ruyigi	Gisuru	Kabuyenge	Control
Ruyigi	Gisuru	Kinama	TNFP

Ruyigi	Gisuru	Mwegereza	T18
Ruyigi	Gisuru	Nkurubuye	Control
Ruyigi	Gisuru	Nyabigozi	T24
Ruyigi	Gisuru	Rwerambere	TNFP
Ruyigi	Gisuru	Taba	Control
Ruyigi	Kinyinya	Bugongo	TNFP
Ruyigi	Kinyinya	Gataba	Control
Ruyigi	Kinyinya	Kabanga	TNFP
Ruyigi	Kinyinya	Kinyinya	TNFP
Ruyigi	Kinyinya	Nyakibere	Control
Ruyigi	Kinyinya	Nyamusasa	T18
Ruyigi	Kinyinya	Vumwe	Control
Ruyigi	Nyabitsinda	Nyakiyonga	Control
Ruyigi	Nyabitsinda	Nyarumuri	T18
Ruyigi	Nyabitsinda	Remba	T18
Ruyigi	Ruyigi	Bunogera	Control
Ruyigi	Ruyigi	Buruhukiro	T18
Ruyigi	Ruyigi	Rutonganikwa	T18
Ruyigi	Ruyigi	Ruyigi rural	TNFP

[a] T24: 15 *collines* assigned to the intervention arm receiving the full *Tubaramure* program from pregnancy to 24 months
T18: 15 *collines* assigned to the intervention arm receiving the full *Tubaramure* program from pregnancy to 18 months
TNFP: 15 *collines* assigned to the intervention arm receiving the full *Tubaramure* program from pregnancy to 24 months, without food rations during pregnancy
Control: 15 *collines* assigned to the control arm of the research study.
More details are provided in the text.

Appendix 3. List of Motor Milestones

Number	Milestone
1	Sit up and hold head straight
2	Lift head and chest when lying on the belly
3	Turn from the belly toward the back
4	Keep head straight, without wavering, when held sitting
5	Move with the belly on the floor (just like swimming)
6	Sit supported by someone or something
7	Sit on his/her own
8	Lift his/her belly when lying on the stomach, and hold on with hands, feet, or knees
9	While lying on the belly, head and chest up, can move using arms and legs
10	Crawl on all fours
11	Stand leaning on someone or something
12	Walk when both hands are held
13	Walk when one hand is held
14	Stand alone, without help, for a short time
15	Stand alone, without help, for a longer period of time
16	Bend at the waist and stand back up without falling (knees are straight or slightly bent)
17	Take a few steps alone, without help from anyone or anything
18	Run
19	Climb stairs walking on feet (not crawling)
20	Throw a ball (hand lifted to the ear)
21	Go up and down the stairs on feet (not crawling)
22	Kick a ball forward
23	Walk forward following a straight line (can take 10 steps)
24	Jump on both feet (both feet up at the same time)
25	Stand on one foot for several seconds
26	Walk backwards following a straight line (can take 10 steps)
27	Walk on tiptoe for four steps (heels are raised)
28	Jump four times without heels touching the ground
29	Jump on one leg 20 times in a row
30	Skip using alternate legs

Appendix 4. List of Language Milestones

Number	Milestone
1	Make sounds while playing alone
2	Make sounds like da, ba, ga, ka, ma
3	Make sounds like ma-ma, da-da, ba-ba
4	Imitate simple sounds like da, ba, or ma, or repeat them
5	When holding something in his/her hand, will give the item when requested
6	Say one word
7	Say goodbye at the right time waving his/her hand
8	Point and make sounds when he/she wants something
9	Point to a cat or chicken when asked to do it
10	Say three words
11	Point to a person walking if asked to do it
12	Say six words
13	Use pronouns "I" and "you"
14	Constantly ask for names of objects
15	Start a lot of questions with "What," "Where," and "Who"
16	Say a lot of words (20 or more)
17	Use plurals when talking
18	Tell what a knife is used for
19	Say full name (last and first names)
20	Say opposite of word "big"
21	Talk about things that took place in the past (e.g., day before) using the correct conjugation