Expanding Coverage of Severe Acute Malnutrition (SAM) Treatment in Kenya

NCT number NCT06441305

Statistical analysis plan for the impact evaluation of R-SWITCH– Research to Improve the Screening for Wasting and the Identification and Treatment of Wasted Children

Version 1.0 – July 14, 2026

1 Background

Child wasting is the 3rd ranked risk factor of Disability-adjusted life years (DALYs) and death among children under five years of age globally (after low birth weight and short gestation) and is the leading risk factor among children 28 days and older. An estimated 47 million children under five years of age (7% globally) suffer from wasting, and approximately one-third (14.5 million) suffer from severe wasting at any given time (1). In 2019, 16.6 percent of U5 DALYs (78 million) and 17.3 percent of deaths (874,000) among children under 5 years globally were attributable to wasting (2).

Despite the burden and impact of child wasting on morbidity and mortality, only a small proportion of severely wasted children are presently identified and admitted into treatment programs. In 2020, an estimated 4.9 million severely wasted children received treatment, approximately a third of the total burden globally (3). Outside of humanitarian settings, this proportion is even lower (estimated to be ~15%). Given the true burden, treatment coverage is likely to be even lower.

Over the past decade operational research has been conducted to explore ways to improve access and coverage and make treatment simpler, more accessible, and more cost-effective. These programmatic innovations and modifications are commonly called “simplified approaches” (4). Most of these simplifications have to-date been tested in small scale, mostly Non-Governmental Organizations (NGO)-supported pilot projects or efficacy trials, which showed these approaches were likely feasible, effective and safe, but should be further researched at larger scale (5–8). Making these simplified or alternative approaches part of global and national health policies, and a more mainstreamed component of routine health services, requires evidence of their feasibility and impact when delivered at scale through national systems.

The Research to improve the screening for wasting and the identification and treatment of wasted children (R-SWITCH) project, born from a collaboration between the United Nations International Children's Emergency Fund (UNICEF) and the International Food Policy Research Institute (IFPRI, the global research lead) with funding from the Bill and Melinda Gates Foundation, aims to tackle this lack of rigorous evidence. The project conducted in Ethiopia and Kenya identifies and tests innovative approaches to detect and treat wasting at different levels of the system (household, community, facility), thereby increasing the proportion of children wasted who receive timely and appropriate treatment.

In Kenya, a protocol was developed to assess the impact of the R-SWITCH intervention on severe wasting treatment coverage using a randomized controlled trial in Turkana. This Statistical Analysis Plan (SAP) documents all the analyses planned for the R-SWITCH study in Kenya. This includes the definition of the population of study, statistical methods to assess impact, subgroup and sensitivity analysis of the impact assessment. The protocol was registered on clinicaltrials.gov (NCT06441305).

2 Study objectives

The overall objective of the study is to assess the impact on wasting treatment coverage, implementation and cost, of the R-SWITCH package of interventions delivered by Community Health Promoters (CHP), which includes scale-up of family-led MUAC, weight-for-age screening, and real-time follow-up of defaulters, supported by an electronic community health information system.

More specifically, the study aims to:

- Assess the impact on treatment coverage in 6–59-month-old children of the R-SWITCH package of interventions.
- Assess the impact on wasting screening coverage in 6–59-month-old children of the R-SWITCH package of interventions.
- Assess the impact on wasting prevalence in 6–59-month-old children of the R-SWITCH package of interventions.
- Assess the reach, adoption by CHPs, and implementation challenges of the R-SWITCH package.
- Assess the unit cost of the R-SWITCH package per beneficiary child.
- If there is an impact, assess the cost-effectiveness of the R-SWITCH package.

This Statistical analysis plan focus on the impact assessment of the R-SWITCH package.

2.1 Study area and population

The Kenya study is conducted in Turkana County situated in the Northwestern Kenya, in the sub counties of Turkana South, Turkana East, and Aroo. The justification for the selection of the study site is as follows: persistent high GAM rates; presence of partners implementing and testing simplified approaches in the management of wasting; the presence of active CUs; and the presence of motivated CHPs receiving stipend. The County has an estimated total population of 153,843 of children below 5 years old with the highest wasting (32.7%) and severe wasting (4.7%) rates recorded in Turkana South and East.

3 Study methods

3.1 Study design

We designed a two-arm, cluster-randomized, unblinded, community-based effectiveness trial to study the impact of the R-SWITCH community package of interventions on wasting treatment coverage. The evaluation will use a before-after repeated cross-sectional study design (at baseline and after 24 months) to compare two groups of villages with a CHP.

The activities of the control and intervention group are summarized in

Table .

Table 2: Intervention package in intervention and comparison groups

Continuum of care	Control "Standard of care"	SWITCH community package of interventions[1] (in addition to standard of care)
Screening and referral	• Paper registers at all levels of care. Community Health Agent (CHA) supervises 10 CHP (all of them being a community health unit), link with a link HF • Register-based HH registration (census) by CHP every 12 months	• Addition of digital monitoring & supervision by CHA • Digitized registration by CHPs every 6 months • **Digital reminders** to CHP to conduct home visit if a child has never been visited in 3 mo.
	• Some family MUAC of variable and unknown coverage/frequency, national guidance developed.	• **Digital reminder for timely training** to family MUAC of all family members through home visit by CHP when child turns 6 months of age. • **Platform for 2-ways messaging**: biweekly SMS reminders to caregivers through the CHP's app; feedback by caregiver
	• Some passive WHZ screening at Health Facility and outreach during Malezi bora (child health week every 6 months).	• **Community level screening based on weight/WAZ** led by CHP every other month (aligned with Community Action Days), • Digital calculation of WAZ. • CHP and Growth Monitoring and Promotion refer to IMAM if WAZ<-3.
Treatment	• Monthly compilation by CHA of the list of defaulters and non-respondents for transmission to relevant CHP during monthly in-person meeting.	• **At the end of every IMAM consultation**, compilation by IMAM responsible of the list of defaulters and non-respondents, for **immediate transmission by CHA to relevant CHP through digital task** of an instruction for immediate follow-up

The intervention targets children aged 6 to 59 months old and their caregivers of the study area living in the catchment area of a community unit, defined as all villages where a CHP is officially appointed under the supervision of a CHA.

The primary study population, in which the primary outcome (treatment coverage) will be measured, is the population of all children suffering from SAM which includes children currently meeting the SAM criteria, and children initially suffering from SAM who are still on their way to recovery until they successfully complete treatment. Specifically, the population of children suffering from SAM is defined as the population of children aged 6 to 59 months old living in the catchment area of a community unit at the time of baseline data collection AND either:

- currently suffering from SAM defined by WHZ/WLZ <-3 (relative to WHO 2006 reference) or MUAC <115 mm or by the presence of bilateral edema, OR
- recovering from SAM through treatment, defined as cases with MUAC≥115mm and WHZ/WLZ≥-3 and no oedema but still receiving treatment as follow-up for an initial SAM condition on their way to full recovery. The definition of receiving treatment is further detailed in section **Error! Reference source not found.**.

The secondary study population comprises all children 6 to 59 months old living in the catchment area of a community unit in the study area, which is the target population for the R-SWITCH intervention. Indeed, to identify the population of children suffering from SAM (currently and recovering), and because this population is likely to be very small, we will need to conduct a census of all villages of the study area with an appointed CHP, in order to measure inclusion criteria. This includes the measurement of anthropometric outcomes which can be used to assess SAM, MAM and wasting prevalence in the secondary study population of children 6 to 59 months old living in the catchment area of a community unit in the study area.

3.2 Study outcomes

3.2.1 Primary study outcomes

For this study, the definition of SAM, MAM and wasting are defined as follows:

- SAM is defined by WHZ/WLZ <-3 (relative to WHO 2006 reference) or MUAC <115 mm or by the presence of bilateral edema.
- MAM is defined by -3≤ WHZ /WLZ<-2 (relative to WHO reference of 2006) or 115mm≤ MUAC <125 mm.
- Wasting is defined by WHZ/WLZ <-2 (relative to the WHO reference of 2006) or MUAC <125 mm or the presence of bilateral pitting edema.

The **primary outcome** of the study is **SAM treatment coverage** (period prevalence) defined as the number of children suffering from SAM (currently or on their way to recovery from SAM through treatment for an initial SAM condition) who currently receive treatment, over the total number of children suffering from SAM (see box below).

$$Period\ prevalence = \frac{C_{in} + R_{in}}{C_{out} + C_{in} + R_{in}} \times 100$$

C in= SAM cases under treatment

C out= SAM cases not under treatment

R in= recovering SAM cases under treatment defined as cases with MUAC≥115mm and WHZ/WLZ≥-3 and no oedema but still under SAM treatment, or under MAM treatment on their way from SAM status to full recovery.

Children currently suffering from SAM will be considered to receive treatment if they both attended an IMAM consultation in the previous 7 days, or in the previous 15 days if compatible with the application

by the health authorities of the IMAM protocol (as reported by caregiver OR by an IMAM treatment card) AND consumed RUTF at least once in the previous 3 days (as reported by the caregiver AND confirmed by observation by the enumerator of at least 1 full or 2 empty RUTF sachets presented by the caregiver).

Children recovering from SAM through treatment will be considered receiving treatment for an initial SAM condition if they both attended an IMAM consultation in the previous 15 days (as reported by caregiver OR by a consultation card) AND either:

- consumed RUTF at least once in the previous 3 days (as reported by the caregiver AND confirmed by observation by the enumerator of at least one 1 full or 2 empty RUTF sachets presented by the caregiver) OR
- consumed RUSF at least once in the previous 3 days (as reported by the caregiver AND confirmed by observation by the enumerator of at least 1 full or 2 empty RUSF sachets presented by the caregiver) AND the child was previously enrolled for SAM treatment immediately prior to MAM treatment (as confirmed by treatment card OR reported by the mother). This corresponds to receiving MAM treatment as follow-up treatment on the way from SAM status to full recovery.

3.2.2 Secondary study outcomes

Secondary outcomes of the impact evaluation include, **in the primary population**:

- **Point prevalence of SAM treatment coverage** defined as the proportion of children with SAM at the time of the survey that are under treatment (see definition under primary outcome) over the total number of children of the primary study population
- **Period prevalence of wasting treatment coverage** in children 6-59 months of age, defined as the proportion of children suffering or recovering from wasting who currently receive treatment. Wasting will be defined by weight-for-height Z-score (WHZ) <-2 (relative to World Health Organization (WHO) 2006 reference) or MUAC <125 mm or by the presence of bilateral edema. Children recovering from wasting will be considered to receive treatment if they both attended an IMAM consultation in the previous 15 days (as reported by caregiver OR by a consultation card) AND either: consumed RUTF at least once in the previous 3 days (as reported by the caregiver AND confirmed by observation of >= 1 full or 2 empty RUTF sachets) OR [consumed RUSF at least once in the previous 3 days (as reported by the caregiver AND confirmed by observation by the enumerator of >= 1 full or 2 empty RUSF sachets) AND child does not suffer from SAM (as defined above)]
- **Coverage of the IMAM platform** defined as the number of children of the primary study population who attended an IMAM consultation in the previous 15 days (as reported by the caregiver or by an IMAM treatment card) over the total number of children of the primary study population
- **Coverage of the community unit platform** defined as the number of children of the primary study population in contact with a CHP in the 2 months preceding the survey for community care-related matter (as reported by the caregiver) over the total number of children of the primary study population

- **Coverage of wasting screening** defined as the number of children of the primary study population screened in the month preceding the survey with a MUAC tape or weighed on a scale (as reported by the caregiver) over the total number of children of the primary study population
- **Coverage of family MUAC** defined as the number of children of the primary study population screened in the month preceding the survey by a family member using a MUAC tape (as reported by the caregiver) over the total number of children of the primary study population
- **Coverage of family MUAC training** defined as the number of children of the primary study population for which at least one family member was trained to use MUAC tape (as reported by the caregiver) over the total number of children of the primary study population
- **Coverage of MUAC tapes** defined as the number of children of the primary study population for which at least one family member owns a MUAC tape (as reported by the caregiver AND confirmed by observation) over the total number of children of the primary study population
- **Coverage of WAZ screening** defined as the number of children of the primary study population weighed in the 2 months preceding the survey by a CHP or a health staff using a scale (as reported by the caregiver) over the total number of children of the primary study population
- **Mean caregivers' knowledge score** on child feeding, health, SAM, and Water, Sanitation and Hygiene (WASH) in caregivers of children of the primary study population
- **Prevalence of SAM** defined as the number of children of the primary study population with WHZ < -3 or MUAC < 115 mm or the presence of bilateral pitting edema over the total number of children of the primary study population
- **Prevalence of MAM** defined as the number of children of the primary study population with -3≤ WHZ <-2 or 115mm ≤ MUAC <125 mm over the total number of children of the primary study population
- **Prevalence of wasting** defined as the number of children of the primary study population with WHZ < -2 or MUAC < 125 mm or the presence of bilateral pitting edema over the total number of children of the primary study population
- **Prevalence of very low MUAC** defined as the number of children of the primary study population with MUAC < -115 mm over the total number of children of the primary study population
- **Prevalence of very low WHZ** defined as the number of children of the primary study population with WHZ < -3 over the total number of children of the primary study population
- **Prevalence of underweight** defined as the number of children of the primary study population with WAZ < -2 over the total number of children of the primary study population
- **Prevalence of severe underweight** defined as the number of children of the primary study population with WAZ < -3 over the total number of children of the primary study population
- **Prevalence of stunting** in the primary study population defined as the number of children of the secondary study population with length-for-age (LAZ)/height-for-age (HAZ) Z-scores< -2 over the total number of children of the primary study population
- Mean HAZ/LAZ and WAZ scores
- Prevalence of appropriate immunization, defined as the proportion of children aged 6-18 months with SAM or enrolled in SAM OTP who received all recommended immunizations

Secondary outcomes of the impact evaluation include, **in the secondary population**:

- **Prevalence of SAM** defined as the number of children of the secondary study population with WHZ < -3 or MUAC < 115 mm or the presence of bilateral pitting edema over the total number of children of the secondary study population
- **Prevalence of MAM** defined as the number of children of the secondary study population with -3≤ WHZ <-2 or 115mm ≤ MUAC <125 mm over the total number of children of the secondary study population
- **Prevalence of wasting** defined as the number of children of the secondary study population with WHZ < -2 or MUAC < 125 mm or the presence of bilateral pitting edema over the total number of children of the secondary study population
- **Mean WHZ/WLZ and MUAC** in the secondary study population.

3.3 Randomization, sample size and sampling

3.3.1 Randomization

The study area included health facilities catchment areas in the three sub counties: 13 in Aroo (another 3 were excluded prior to randomization because of security issues), 17 in Turkana East and 15 in Turkana South.

We stratified the random allocation of clusters by sub-county. Within each respective sub-county, respectively 12, 14 and 14 health facilities were randomly selected by simple random sampling using Stata's sample command. The selected clusters were then randomly allocated in a 1:1 ratio, with the first half in the randomized list assigned to the R-SWITCH intervention group and the remaining half clusters assigned to the control group. Randomization was conducted by an investigator based in Dakar and with no knowledge of the study area. To ensure the randomness and replicability of this draw, a random seed (the number used to initiate the draw) was obtained from the random number generator on random.org. Of the 40 health facilities catchment areas, 36 covered one community unit each, and four covered two community units each. By chance, the latter were balanced across study groups (two in intervention and two in control). The intervention was implemented in the 20 intervention clusters, including 22 community units. For operational reasons, it was challenging to survey 44 community units at endline. Therefore, in a second stage, for the four facilities with two CUs each, one CU was selected at random prior to participant sampling at endline. All subsequent sampling activities were conducted within the selected CU.

3.3.2 Sample Size

We used the clustersampsi command in Stata (version 18.0, Statacorp LLC, College station, Texas) to calculate the necessary sample sizes for the repeated cross-sectional study.

We assumed a design effect of 2, a type I error of 5% and a statistical power of 90%. Due to poor availability of updated treatment coverage estimates, we considered a treatment coverage of 50% because it is the conservative scenario in terms of sample size (i.e. it requires the largest sample size). We considered that with 43 functioning community units linked to 40 health facilities in the study area, we could use 20 clusters in each treatment arm.

Varying the number of children in the 6-59 months age range from CHPs registration numbers (23,692) and the 2019 census (34,813) and under the lowest (4.06%) and the highest (9.83%) SAM prevalences observed in the SMART surveys between 1017 and 2022, we assume that with a screening that reaches at least 85% of the SAM children, we will enroll between 818 and 2,909 SAM children. Enrolling in the study all children with SAM identified in 40 community units would therefore allow to detect an increase in the severe wasting treatment coverage of 9-16 percentage points.

We estimated this detectable difference to be acceptable (i.e. the intervention is likely to lead to an increase in treatment coverage of at least 16 pp).

For secondary outcomes measured in the primary study population, we used the same assumptions of 40 clusters, 818-2909 enrolled children, design effect of 2, type I error of 5% and statistical power of 90%, to calculate detectable differences as presented in Table 1.

Table 1: Detectable differences for secondary outcomes in the primary study population.

Secondary outcome	Baseline prevalence or mean ± SD	Detectable difference
Point prevalence of SAM treatment coverage	50% (conservative scenario)	9-16 pp
Period prevalence of wasting treatment coverage	50% (conservative scenario)	9-16 pp
Coverage of the IMAM platform	50% (conservative scenario)	9-16 pp
Coverage of the community unit platform	50% (conservative scenario)	9-16 pp
Coverage of wasting screening	50% (conservative scenario)	9-16 pp
Coverage of family MUAC	50% (conservative scenario)	9-16 pp
Coverage of family MUAC training	50% (conservative scenario)	9-16 pp
Coverage of MUAC tapes	50% (conservative scenario)	9-16 pp
Coverage of WAZ screening	50% (conservative scenario)	9-16 pp
Mean caregivers' knowledge score	38 ± 16 points	5.3 points
Prevalence of SAM	50% (conservative scenario)	9-16 pp
Prevalence of MAM	50% (conservative scenario)	9-16 pp
Prevalence of wasting	50% (conservative scenario)	9-16 pp
Prevalence of very low MUAC	50% (conservative scenario)	9-16 pp
Prevalence of very low WHZ	50% (conservative scenario)	9-16 pp
Prevalence of underweight	50% (conservative scenario)	9-16 pp
Prevalence of severe underweight	50% (conservative scenario)	9-16 pp
Prevalence of stunting	50% (conservative scenario)	9-16 pp
Mean HAZ/LAZ	-0.9 ± 1.1 Z-scores (DHS 2022)	0.37 Z-scores
Mean WAZ	-0.6 ± 1.1 Z-scores (DHS 2022)	0.37 Z-scores
Prevalence of appropriate immunization	50% (conservative scenario)	9-16 pp

For secondary outcomes measured in the primary study population, we used the assumptions of 40 clusters, reach of 85% of the 23,692 children followed by CHP, design effect of 2, type I error of 5% and statistical power of 90%, to calculate detectable differences as presented in Table 2.

Table 2: Detectable differences for secondary outcomes in the secondary study population.

Secondary outcome	Baseline prevalence or mean ± SD	Detectable difference
Prevalence of SAM	4.1-9.8 %	1-2 pp
Prevalence of MAM	15%	2 pp
Prevalence of wasting	19-25 %	3 pp
Mean WHZ/WLZ	-0.2 ± 1.1 Z-scores	0.07 Z-scores
Mean MUAC	136 ± 52 mm	3.45 mm

3.3.3 Sampling and study inclusion criteria

This study will not involve sampling of subjects. Indeed, to survey a sufficient sample size, we will conduct a complete enumeration of the entire secondary study population to determine eligibility for the survey (screening campaign). Door-to-door visit to households will be conducted (with support from village gatekeepers) to invite all children aged 6-59 months of age to get screened for wasting through the measurement of MUAC, length (below 24 months of age) or height (above 24 months of age) and weight, and assessment of the presence of bilateral pitting oedema. SAM treatment status will be investigated through the caregiver (IMAM visit consultation in the past 15 days and consumption of treatment product, RUTF or RUSF).

All children meeting the inclusion criteria for the primary study population will be enrolled in the household and child survey.

The **criteria for inclusion of** children in both the primary and secondary study populations are:

- Household in a village of the study area covered by a CHP (although the child may or may not be registered by a CHP) AND
- Child is 6-59.9 months of age AND
- Caregiver consents to be part of the study.

In addition, children meeting all eligibility criteria above will also be **eligible to be part of the primary study population** if they meet at least one of the following **inclusion criteria**:

- WHZ/WLZ < -3 (relative to WHO 2006 reference) OR
- MUAC <115 mm OR
- Presence of bilateral edema OR
- receiving treatment as follow-up for an initial SAM condition on the way to full recovery (see definition in section **Error! Reference source not found.**).

For both populations, the **exclusion criteria** is:

- Congenital malformation that makes anthropometric measurements impossible.
- Length is below 54 cm or height is above 120cm.

The last exclusion criteria will exclude children which are extremely unlikely to be in the eligible age range.

4 Measurements and indicators compilation

Age will be calculated in days counting the number of days between the date of birth and the date of data collection. Date of birth will be collected from birth certificate, mother-child booklet or other official document (birth notification, etc.). When no written record is available, date of birth will be estimated by the enumerator using a detailed calendar of local events.

All weight, length and MUAC measurements will be taken in duplicate by a trained enumerator and a trained assistant, and a third measurement will be taken if the difference between the two repeated measurements is >5mm for length and MUAC and >300g for weight. The average of the two closest repeated measurements will be used for the calculation of nutritional status indicators.

The child weight-for-height and height-for-age z-scores will be calculated using the *zscore06* command in Stata (9), which uses the World Health Organization's growth standard ((10).

MUAC screening will be defined as caregiver recall that MUAC was measured in the target child.

SAM treatment coverage will be calculated at individual level on the sub-group of children with wasting and children recovering from wasting, using the formula:

$$Period\ prevalence = \frac{C_{in} + R_{in}}{C_{out} + C_{in} + R_{in}} \times 100$$

where Cin = current SAM cases under treatment, Cout = current SAM cases not under treatment, Rin= recovering SAM cases under treatment. Children with wasting and recovering children will be identified using anthropometric measures.

Principal component analysis will be used to generate a household wealth status score (11). We will include household-level data on livestock owned, dwelling occupation status, household' characteristics (floor material, number of rooms, etc.), the main source of electricity, and assets owned by 5 to 95% of the study sample. Tertiles of loadings on the principal component with the highest eigenvalue will be used.

5 Data management

For survey work, interviewers will use Android tablets with specialized software for computer interviewing (Survey Solutions software). Immediately after validation of the questionnaire by the interviewer, each questionnaire will be synchronized via an internet connection on a secure server, and the questionnaire will be automatically deleted from the interviewer's tablet. The server will be accessible only to IFPRI and Kenyatta University investigators. After checking the quality of the data collected, the questionnaire can be returned to the interviewer's tablet for field corrections. Once these corrections have been made, the same synchronization process will allow the data to be stored on the secure server.

Management of all research data will follow IFPRI's institutional protocol for research data management, overseen by the IFPRI Data Governance Committee, as well as UNICEF data governance standards. Data will be stored on a secure server accessible only to principal investigators. All data will then be carefully anonymized and de-identified so that the privacy of participants and research subjects is fully protected.

Databases will be anonymized using identification codes. No names, GPS coordinates, dates of birth or other identifying data will be stored in the databases.

At the time of publication of scientific articles presenting primary results, the fully anonymized databases (quantitative data) will become a public good and will be made available to the scientific community, government, and partners.

The raw qualitative data which cannot be fully de-anonymized will remain exclusively stored on the secure server until being destroyed once analyzed.

6 Statistical analysis

6.1 Data sources

The data collected between June and August 2024 will be used to check for baseline balance between the two study groups regarding sample characteristics, and for adjustment for the baseline cluster-level mean of the outcome and some strong prognostic factors at baseline. Endline data collected from June and July 2026 will be used to assess the impact of the intervention and to adjust for some individual-level factors that are pre-determined (pre-intervention) or not modifiable by the intervention, and to check for baseline balance between the two study groups at endline on factors that are pre-determined (pre-intervention) or not modifiable by the intervention.

6.2 Treatment of missing data for primary study outcomes

We will first assess and report the extent of the missing data for both baseline and endline. No anthropometric imputations will be performed to determine SAM status. Children with insufficient information to determine eligibility (SAM status and treatment status) will be excluded from the analysis.

If the missing data on the primary outcome is less than 5%, we will conduct complete-case analysis. If there is more than 5% missing data, a sensitivity analysis will be performed to calculate treatment coverage in the best-case scenario (all missing data are receiving treatment in the intervention group and not receiving treatment in the control group) and worst-case scenario (all missing data are not receiving treatment in the intervention group and receiving treatment in the control group). Missing covariates will be imputed using the cluster mean (12).

6.3 Statistical analysis methods

Data will be analyzed on intent-to-treat basis. Data management, data cleaning, and statistical analyses will be done using Stata 19.0 (Statacorp, USA). The statistical significance will be set at 5% for main effects and 10% for interactions. All statistical tests will be two-sided.

Intervention effects after 2 years of program implementation will be estimated using mixed-effects regression models following an ANCOVA approach. Using an ANCOVA approach improves precision and power compared with analyzing endline outcomes alone.

For continuous outcomes, mean differences will be estimated using linear mixed-effects ANCOVA models with robust estimation of standard errors. For binary outcomes, risk differences will be estimated using linear probability mixed-effects ANCOVA models with robust estimation of standard errors.

The results will be from minimally adjusted and fully adjusted models. The minimally adjusted ANCOVA model is justified because randomization is expected to balance both measured and unmeasured confounders, and will only use community unit as a random intercept and sub-county stratum (used to restrict randomization) as fixed effect, to reflect the study design.

Fully adjusted ANCOVA models can help improve precision and power when adjusting for a prespecified set of covariates thought to be prognostic of the outcome; and could potentially reduce bias when adjusting for covariates prognostic of differential recruitment or outcome missingness(12,13). Therefore, fully adjusted models will additionally include as fixed effects the baseline cluster-level value of the outcome, other known determinants of the outcome not modifiable by the intervention (see details below), as well as a covariate indicating if the household is nomadic, which is frequent in the study area. Nomadic households may move in or out of an intervention area due to their knowledge of the intervention being delivered, which can create differential recruitment.

Fully adjusted models will be the primary models for primary and secondary outcomes, and minimally adjusted models will also be presented for transparency, as recommended (14).

Covariates thought to be prognostic of the outcome are informed by an analysis of determinants of treatment outcomes conducted on baseline data. Household- or caregiver-level determinants associated with higher likelihood of being under treatment included greater household wealth, greater maternal knowledge on child health, greater maternal decision power score and lower maternal depression symptoms score. Child-level determinants of treatment coverage included younger age, being a girl and lower height-for-age Z-score. Services and practices which increased treatment coverage included both the measure of MUAC by a household member in the past month and self-referral for care seeking, as well as referral from a (community) health staff at the health facility, or during outreach health services provided by health center staff in the village (when children may be seeking for a range of other health services). On the contrary, lower treatment coverage was more likely in case of recent history of missing a treatment visit (while enrolled into a current or previous treatment program) or in the rare cases (<3%) of treatment product stockout experienced during a treatment. Most of these factors where intended to be modified by the intervention and should therefore not be adjusted for. However, the intervention was unlikely to modify household wealth, child age and child sex. These three covariates measured at individual level at endline will be included in impact models on the primary outcome and on other treatment coverage secondary outcomes.

For impact models on anthropometric outcomes, covariates will include child age, child sex and primiparity measured at individual level at endline, since these variables are known to be prognostic factors for child nutritional status.

6.4 Considerations for analysis

6.4.1 Baseline characteristics

The baseline characteristics of the subjects will be presented by study group. Following the CONSORT guidelines, we will not statistically test for differences between study groups at baseline (15). Imbalance will be considered for absolute differences >10 pp (binary outcomes) or >10 % (continuous outcomes).

The presentation of baseline characteristics will be done as follows:

- Categorical variables: frequencies and percentages, as appropriate. Percentages will be calculated based on the number of participants for whom data is available.
- Continuous variables: mean and SD or median and interquartile range, as appropriate.

For characteristics that may lie on the causal pathway between the intervention and outcome, only characteristics measured at baseline will be presented. For characteristics determined prior to intervention exposure or very unlikely to be affected by the intervention, we will also present balance in the endline sample. The list of these characteristics was pre-specified before analysis and is detailed in the next session.

Characteristics	**Comments/References**	**Round presented**
Village level		
Distance to the health facility		Baseline
Household level		
HH size (mean/SD)		Baseline and Endline
Number of under 5 children in the household (mean/SD)		Baseline and Endline
Dependency ratio (%)	Number of individuals aged 0–14 years and 65 years and above divided by the number of individuals aged 15–64 years, expressed per 100 working-age individuals, following standard definitions from the United Nations and World Bank.	Baseline and Endline
Nomadic HH (%)		Baseline and Endline
Polygamous HH (%)		Baseline and Endline
HH food insecurity (mean and proportion of food secure HH)	Coates J, Swindale A, Bilinsky P. Household Food Insecurity Access Scale (HFIAS) for measurement of food access: indicator guide. Washington, DC: FANTA; 2007. https://doi.org/10.1007/s13398-014-0173-7.2	Baseline
SES status classification (tertiles, %)	Principal component analysis will be used to construct a proxy household Socio-Economic Score (SES) using ownership of various assets (an asset is excluded if ownership is below 5% or above 95% on HH level), housing materials, primary source of lighting, primary energy source, home ownership and livestock ownership. Tertiles of the first principal component (i.e., the one with the highest eigenvalue) will be used.	Baseline
Improved water treatment technologies used (%)	USAID (2021) Water and development- Indicator handbook. Washington DC	Baseline

Characteristics	Comments/References	Round presented
Improved sanitation facility (%)	USAID (2021) Water and development- Indicator handbook. Washington DC	Baseline
Handwashing station with soap available (%)	USAID (2021) Water and development- Indicator handbook. Washington DC	Baseline
Improved primary water source (%)	USAID (2021) Water and development- Indicator handbook. Washington DC	Baseline
Male head of household (%)		Baseline and Endline
Paternal level		
Father is member of the HH (%)		Baseline and Endline
Maternal level		
Biological mother (%)		Baseline and Endline
Age (mean/SD)		Baseline and Endline
Marital status (%)		Baseline and Endline
Main language (%)		Baseline and Endline
Main religion (%)		Baseline and Endline
Mother school attendance (%)		Baseline and Endline
Has an income generating activity (%)		Baseline
Digital literacy: Access to a phone (%)		Baseline
Digital literacy: language of sending/receiving messages and vocals (%)		Baseline
Knowledge score in 4 domains (breastfeeding, hygiene, health, supplements)	Self-developed questionnaire of 23 questions based on training curricula of CHPs. Mean scores are presented by domain.	Baseline
Possible depression (EPDS≥10)	Cox JL, Chapman G, Murray D, Jones P. Validation of the Edinburgh postnatal depression scale (EPDS) in non-postnatal women. J Affect Disord; 1996;39:185–9.	Baseline
Maternal MUAC (mm) (mean/SD)		Baseline
Mother weight (kg) (mean/SD)		Baseline
Status of women in decision making, community, mobility	Self-developed questionnaire inspired by Pro-WEAI questionnaire to measure women power over decision making about household diet, health	Baseline

Characteristics	Comments/References	Round presented
	decision for the child, community involvement and mobility (women has to inform or ask for permission). No scores is calculated.	
Mother perceived lack of time	A set of 5 questions to measure mother's self-appreciation of time for leisure and workload.	Baseline
Assisted delivery of last pregnancy (%)		Baseline
Child		
Age (mean/SD)		Baseline and Endline
Male (%)		Baseline and Endline
Primiparity (%)		Baseline and Endline
Weight-for-length Z-score (mean/SD)	Relative to WHO 2006 growth reference	Baseline
MUAC (mean/SD)		Baseline
Wasting (WLZ<-2 or MUAC<125mm or edema) (%)		Baseline
Severe Acute malnutrition (WLZ<-3 or MUAC<115mm or edema) (%)		Baseline
Length-for-age Z-score (mean/SD)	Relative to WHO 2006 growth reference	Baseline
Stunting (LAZ<-2) (%)		Baseline
Weight-for-age Z-score (mean/SD)	Relative to WHO 2006 growth reference	Baseline
Underweight (WAZ<-2) (%)		Baseline
Severe underweight (WAZ<-3) (%)		Baseline
Treatment status (%)	Based on the definition of being under treatment developed in our protocol (section primary study outcomes).	Baseline
Screening status (%)	Percentages of children screened by MUAC, weight, length in the last 30 days, 60 days and 90 days according to mother's recall.	Baseline
Immunization status (%)	Vaccines received since birth according to observation of the vaccination card/booklet	Baseline

6.4.2 Sensitivity analyses

1. Sensitivity analyses will assess whether findings may be impacted by migration-related contamination. Primary analyses will be repeated among children residing in the study cluster for

at least 12 months prior to the endline survey (or since birth for children younger than 12 months), as measured at endline.

2. As explained above, if there is more than 5% missing data, a sensitivity analysis will be performed to calculate treatment coverage in the best-case scenario (all missing data are receiving treatment in the intervention group and not receiving treatment in the control group) and worst-case scenario (all missing data are not receiving treatment in the intervention group and receiving treatment in the control group).
3. For rare outcomes (<5%), such as SAM prevalence, or if the sample size of some clusters is small at baseline (<=5), sensitivity analyses will be conducted by rerunning the fully adjusted models without adjustment for baseline cluster-level prevalence. When outcomes are rare and/or cluster sample sizes are small, baseline prevalence estimates may be subject to substantial sampling variability and may therefore contain limited information on true underlying differences between clusters.
4. We will assess whether impacts on the primary outcome are sensitive to the definition of wasting, by assessing impact on SAM treatment coverage when SAM is defined with 2 alternative definitions: 1) MUAC < 115 mm or oedema, and 2) WLZ < -3 Z-scores or oedema.
5. We will assess the sensitivity of the estimated intervention effects on the primary outcome to the definition of treatment receipt. In the primary analysis, a child reported to have consumed treatment within the previous 3 days will be considered to have verified treatment receipt only if the caregiver can provide supporting evidence to the enumerator by showing at least one full RUTF/RUSF sachet or two empty sachets. As a sensitivity analysis, we will use a broader definition of verified treatment receipt, whereby evidence recorded in the mother–child booklet that the child attended a SAM treatment consultation within the previous 7 days or a MAM treatment consultation within the previous 14 days will be considered equivalent to showing at least one full RUTF/RUSF sachet or two empty sachets.

6.4.3 Secondary analyses of effect modification

Studying effect modification of an intervention offers important insights into the mechanism of a possible impact. Effect modification analyses are exploratory and are intended to assess whether intervention effectiveness differs across contextual, household, caregiver, child, and health-system characteristics that may influence exposure to intervention activities, care-seeking behavior, or access to SAM treatment services. Potential effect modifiers include measures of health service accessibility, community health worker capacity, caregiver resources and autonomy, household vulnerability, and child characteristics.

Given that these tests are exploratory, we will assess interaction terms of pre-determined potential modifiers with study group allocation at a statistical significance set at 10%. Statistical models will be the fully adjusted models as described in earlier sections, adding the main effect of the potential modifier and its interaction term with the study group. If there are indications of effect modification by one or more covariates (P-value>0.10), we will present impact results stratified by covariate level.

For subgroup analyses involving a small number of clusters (<15 clusters), where feasible, small-sample corrections to the degrees of freedom (Kenward–Roger) will be used rather than robust estimation of standard error to improve the accuracy of confidence intervals and hypothesis tests. Results will be

interpreted cautiously because standard error and confidence interval estimation may be less reliable in small samples.

As recommended, potential modifiers are pre-specified below, and the hypothesized effect modification is justified (16).

Potential modifier (measured at endline)	Hypothesized effect modification
Sub-county	The intervention effect may vary across sub-counties due to differences in health service availability, implementation fidelity, security conditions, geography, and baseline treatment coverage.
CHP is volunteer	The intervention may be less effective where CHPs are volunteers due to competing livelihood demands and reduced incentive.
CHP is literate in Kiswahili or English	The intervention may be more effective when CHPs are literate because they can better use the eCHIS (available in English or Kiswahili), understand training materials, record information accurately, and communicate with health services.
CHP is equipped with working phone with eCHIS	The intervention may be more effective where CHPs have functioning digital tools that facilitate screening, tracking, referral, and follow-up.
Conflict affected village (highest tertile)	The intervention effect may be reduced and/or noisier in highly conflict-affected villages because insecurity can disrupt health service access, community outreach, and follow-up activities, and lead to migration.
Conflict affected household (highest tertile)	The intervention may be less effective among households directly affected by conflict because displacement, insecurity, and livelihood disruptions can impede access to treatment services.
Nomadic household	The intervention may be less effective and/or noisier among nomadic households because frequent movements of some or all household members can reduce exposure to community-based activities and continuity of care, or facilitate cross-contamination of the household between study groups.
Distance to health facility (village)	The intervention effect may be attenuated among households located farther from health facilities because transportation and travel burdens may limit treatment uptake following referral.
Time needed to go to health facility	The intervention effect may be smaller among households requiring longer travel times because greater opportunity costs and transportation barriers may reduce treatment uptake.
Household wealth status	The intervention may have greater impact among wealthier households, which face fewer financial barriers to accessing care.

Potential modifier (measured at endline)	Hypothesized effect modification
Father is member of the HH	The intervention may be more effective when fathers are present if this increases financial, logistical, or decision-making support for seeking treatment.
Main caregiver is the biological mother	The intervention may be more effective when the biological mother is the primary caregiver because she may be more consistently exposed to intervention activities and more involved in health-related decisions.
Maternal access to a phone	The intervention may be more effective among mothers with phone access because they can receive reminders, communicate with CHPs, and seek care more easily.
Maternal literacy in Turkana, Kiswahili or English	The intervention may have greater impact among literate mothers who can better understand reminders messages and referral instructions.
Maternal knowledge on health and supplements	The intervention may have greater impact among mothers with higher knowledge because they may be better able to understand the consequences of malnutrition, recognize illness symptoms, understand the importance of treatment, act on referrals.
Maternal possible depression (EPDS≥10)	The intervention may be less effective among mothers experiencing depressive symptoms, which may reduce care-seeking, adherence to referrals, or engagement with health services.
Maternal Decision making power	The intervention may have greater impact among women with higher decision-making autonomy because they may be better able to act on referrals and seek treatment for their child.
Mother connected with community	The intervention may be more effective among mothers with stronger community connectedness due to greater exposure to information, peer support, and community health activities.
Maternal mobility	The intervention may have greater impact among women with higher mobility, as they may be better able to independently seek care, attend health facilities, follow referrals, and engage with community health activities without requiring permission from others.
Maternal perceived lack of time	The intervention may have smaller effects among mothers reporting time constraints, as competing responsibilities may limit their ability to seek treatment or attend follow-up visits.
Perception of stigma	The intervention may have reduced effectiveness among caregivers who perceive greater stigma related to malnutrition or treatment services, potentially discouraging care-seeking.

Potential modifier (measured at endline)	Hypothesized effect modification
Absence of mother in past month	The intervention may be less effective among children whose mothers were recently absent because reduced caregiver availability may hinder screening, referral uptake, and treatment seeking.
Primiparity	The intervention may be more effective among first-time mothers if they are more receptive to health information, or less effective if they have less caregiving experience.
Child age below 24 mo	The intervention effect may differ by child age because younger children may have more frequent contact with health services and caregivers may be more responsive to health advice during this period.
Child sex	The intervention effect may differ by child sex due to gender-related differences in caregiving practices or caregiver perceptions of illness and vulnerability.

7 References

1. UNICEF, WHO, World Bank Group. Levels and Trends in Child Malnutrition: Key Findings of the 2020 Edition of the Joint Malnutrition Estimates. 2020.

2. UNICEF, WHO, World Bank Group. Levels & Trends in Child Mortality: Report 2020 - Estimates Developed by the UN Inter-Agency for Child Mortality Estimation. 2020.

3. WHO. Global Action Plan on Child Wasting: A framework for action to accelerate progress in preventing and managing child wasting and the achievement of the Sustainable Development Goals. 2019.

4. UNICEF. Treatment of Wasting Using Simplified Approaches - A Rapid Review [Internet]. [cited 2023 Nov 28]. Available from: https://www.unicef.org/documents/rapid-review-treatment-wasting-using-simplified-approaches

5. Bailey J, Opondo C, Lelijveld N, Marron B, Onyo P, Musyoki EN, Adongo SW, Manary M, Briend A, Kerac M. A simplified, combined protocol versus standard treatment for acute malnutrition in children 6–59 months (ComPAS trial): A cluster-randomized controlled non-inferiority trial in Kenya and South Sudan. PLoS Med 2020;17:e1003192.

6. Marron B, Onyo P, Musyoki EN, Adongo SW, Bailey J. ComPAS trial in South Sudan and Kenya: Headline findings and experiences. 2019 Jan 7;(60):19. Field Exch 2019;19–22.

7. Kimani-Murage EW, Pythagore H, Mwaniki E, Daniel T, Samburu B, Cuellar PC, Mbochi R, Njiru J, Wangare L, Karimurio L, et al. Integrated and simplified approaches to community management of acute malnutrition in rural Kenya: a cluster randomized trial protocol. BMC Public Health 2019;19:1253.

8. Grant A, Njiru J, Okoth E, Awino I, Briend A, Murage S, Abdirahman S, Myatt M. Comparing performance of mothers using simplified mid-upper arm circumference (MUAC) classification devices with an improved MUAC insertion tape in Isiolo County, Kenya. Archives of Public Health [Internet] 2018;76:11. Available from: https://archpublichealth.biomedcentral.com/articles/10.1186/s13690-018-0260-x

9. Leroy JL. ZSCORE06: Stata command for the calculation of anthropometric z-scores using the 2006 WHO child growth standards. 2011.

10. World Health Organization. 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; 2006.

11. Vyas S, Kumaranayake L. Constructing socio-economic status indices: How to use principal components analysis. Health Policy Plan 2006;21:459–68.

12. Hemming K, Hall J, Copas A, Watson S, Taljaard M, Hooper R. Covariate adjustment in cluster randomised trials: a practical guide. BMJ [Internet] British Medical Journal Publishing Group; 2025 [cited 2026 Jul 9];391. Available from: https://www.bmj.com/content/391/bmj-2025-084194

13. Hooper R, Forbes A, Hemming K, Takeda A, Beresford L. Analysis of cluster randomised trials with an assessment of outcome at baseline. BMJ [Internet] 2018 [cited 2026 Jul 9];360:1121. Available from: http://dx.doi.org/10.1136/bmj.k1121

14. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. BMJ : British Medical Journal [Internet] BMJ Publishing Group; 2004 [cited 2026 Jul 9];328:702. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC381234/

15. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. Bmj 2012;345:e5661–e5661.

16. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet [Internet] Elsevier B.V.; 2000 [cited 2026 Jul 9];355:1064–9. Available from: https://www.thelancet.com/action/showFullText?pii=S0140673600020390