Statistical analysis plan for:

IRAM Mali – Integrated Research on Acute Malnutrition: a clustered randomized controlled trial in Mali

Version 1.0: August 24, 2021
Version 2.0: September 22, 2022

Summary of changes between v1.0 and v2.0:

1. The loss of 3 study clusters because of serious insecurity threats reduced the sample from 45 health center catchment areas to 42.
2. The sponsor agreed to extending the study with two months.
3. Sample sizes were recalculated reducing the number of clusters from 45 to 42 and adding two months of longitudinal follow-up

Study registration on clinicaltrials.gov (NCT04872088) on May 4, 2021

Background

Worldwide, 47 million children under five years of age suffer from wasting [1]. Wasting significantly increases the risk of death: it kills 875,000 children under the age of five per year. All Member States of the World Health Assembly (WHA) have agreed to reduce and maintain the prevalence of wasting to less than 5% by 2025[2]. However, the prevalence of wasting remains persistently high in many West African countries such as Mali (14.8%), Mauritania (13.5%), Chad (12.9%) and Niger (10.1%). Wasting trends suggest that the WHA goal will not be achieved in these countries, despite their commitment to addressing wasting as expressed in their nutrition policies.

There is an urgent need to test innovative solutions to prevent wasting and increase screening, adoption and adherence to treatment. Options include integrating prevention into screening to increase coverage and prevent wasting; strengthening referral processes for wasted children and supporting and encouraging parents to enroll their children in treatment; establishing outreach screening and treatment units; transferring some treatment responsibilities to community health workers; and/or simplifying treatment procedures for both parents and service providers.

In response to this urgent need, UNICEF and IFPRI started a multi-country partnership to generate evidence on wasting prevention and treatment interventions in four countries: Chad, Mali, Mauritania, and Niger. Several interventions aiming at addressing gaps along the spectrum of wasting will be integrated and implemented at community and health center levels (first and second lines) given their demonstrated importance in the 4 countries of interest.

In Mali, an impact evaluation will be conducted to generate strong evidence on the effectiveness of these integrated interventions to reduce wasting among young children.

[1] Defined in children aged 0-59 months as a weight-for-height score of less than -2 z-score according to WHO weight-for-height reference; depending on the context and the measurement tools used, wasting is also defined by a brachial perimeter of less than 125 mm in a child aged 6-59 months or by the presence of bilateral edema (0-59 months).

[2] Resolution WHA65.6. Comprehensive Implementation Plan on Maternal, Infant and Young Child Nutrition." 2012. In *Sixty-Fifth World Health Assembly, Geneva, 21–26 May 2012*, Geneva.

Study objectives

The overall objective of the study is to assess the impact of an integrated package of interventions covering prevention, early detection, treatment, and relapse prevention on the prevalence of child wasting.

The implementation of these interventions will be led by World Vision Mali in collaboration with the health services of the Koutiala health district (Sikasso region, Mali), and will take place at the health center and community levels, and include i) a prevention component combining the strengthening of Nutrition Activity Support Groups (NASG) (which will conduct home visits to deliver behavioral change communications and cooking demonstrations) and the distribution of a nutritional supplement (Small-quantity lipid-based nutrient supplements) to children over 6 months of age; ii) a component related to strengthening screening and referral that will involve families (MUAC family approach) and community groups; iii) a treatment component that includes strengthening the national CMAM protocol currently in use in Mali and supporting community groups through home follow-up visits of wasting cases; and iv) a final relapse prevention component combining home follow-up visits by NASGs of children that recovered from wasting and the distribution of a preventive nutritional supplement.

Study methods

Study area and population

The study will be conducted in the Koutiala health district located in Sikasso (region of Mali), which includes 45 community health centers (called CSCOM) and 1 referral health center (called CSRef). Population statistics for this district show a total of 145,845 children aged 6-59 months in 2019, with 6.5% of them suffering from wasting [6].

Study design

Two study groups will be formed, a control and an intervention group receiving the integrated package of interventions.

The activities of the control and intervention group are summarized in Table 1.

Table 1: Description of Prevention, Screening, Treatment, and Relapse Prevention Activities in the Control and Intervention Groups

	Mali	
	Control*	Intervention
Prevention		
Usual NASG activities (Home visits and group counselling on nutrition and health)	X	X
Integrated NASG (or intensive)		
- NASG created proportionate to population size of village		X
- Frequent home visits (once a month)		X
- Strengthening of group counselling / cooking demonstrations activities		X
- Distribution of SQ-LNS to 6–17-month children		X
Screening		
Usual activities of the NASGs (screening using MUAC criterion during home visits and group counselling)	X	X
Family MUAC		X
Screening with Weight-for-height / MUAC during the 7th month of the IRAM project	X	X
Treatment		
National CMAM Program (CSCom and CHW Site)	X	X
Integrated NASG (or intensive)		

Follow-up of cases under treatment by intensive home visits (twice a month)	X
Relapse prevention	
Integrated NASG (or intensive)	
-Intensive home visits (twice a month)	X
-Distribution of nutritional supplements to 6–24-month children post-treatment	X

**In Koutiala villages, NASGs were created by the health district and UNICEF in 2019. However, their monitoring was stopped in 2019. It is not known at this time whether the NASGs are still operational.*

The study was designed as a two-arm, cluster-randomized, nonblinded effectiveness study. A cluster-randomized design was used because individual randomization of the NASG intervention was not feasible. A cluster was defined as the catchment area of a first-line health center. All health center catchment areas included in the study were first grouped into 4 strata according to their characteristics to obtain a more balanced randomization result.

A population-representative longitudinal study design (Cohort 1 or Main cohort) was used enrolling children aged 6 to 6.9 months from May 2021 until November 2021. In addition, a second cohort (cohort 2 or Treatment cohort) was constituted consisting of all children admitted to MAM OTP (French: URENAM) and SAM OTP (French: URENAS) provided by first-line health centers and community health workers from May 2021 until February 2022. Finally, study field workers home visited a subsample of children three months after their discharge from SAM or MAM OTP to assess the impact of the intervention on post-treatment relapse rates (cohort 3 or Relapse prevention study). While the Main cohort study is an experimental study[3], the Treatment cohort and Relapse prevention study are quasi-experimental in nature. The quasi-experimental character of these twee sub-studies implies that the random character of the intervention and comparison group is not fully guaranteed. More precisely, children enrolled in the Treatment cohort may not be comparable between study groups because the BCC and SQ-LNS may prevent a proportion of the usual prevalence of wasting. Similarly, a higher screening coverage or screening by family MUAC may refer children with wasting with different characteristics as compared to those in the comparison group. Therefore, any impact or lack of impact of the NASG intervention to support the OTP services needs to be interpreted with caution, as it may be a result of both the intervention and any different characteristics of the enrolled samples of children. For the same reasons, the results of the relapse prevention study must be interpreted bearing in mind possible impacts of the preceding preventive BCC and SQ-LNS and improved screening and referral (studied in the main cohort), as well as possible impacts on the adherence and recovery rates of MAM and SAM OTP (studied in the treatment cohort).

Study outcomes

Primary study outcomes

1. The longitudinal prevalence of wasting in children enrolled at the age of 6 months followed monthly until the age of 12 months or until the end of the study, whatever comes first (cohort 1 or Main cohort). This indicator is defined for each child as the number of visits during which nutritional wasting is observed divided by the total number of monthly visits made (by the interviewers).
2. The recovery rate in children enrolled at the age of 6-23 months for a treatment of up to 3 months in URENAM and URENAS treatment services and followed until their discharge (cohort 2 or

[3] The main cohort study is an experimental cluster-randomized controlled study that guarantees (if random allocation of clusters is successful) that the samples of children enrolled in the intervention and comparison group have similar household, caregiver, and child characteristics. Therefore, any observed impacts can be exclusively associated with the intervention.

Treatment cohort). This indicator is defined as the number of children declared recovered from wasting according to the criteria of the national program (WLZ-score≥-1.5 or MUAC≥ 125mm, and absence of bilateral edema during two consecutive visits, within 12 weeks after enrollment in the program) divided by the total number of enrolled children.

3. The prevalence of relapse after URENAS or URENAM treatment defined as the proportion of children (9-18 months of age) with WLZ-score <-2 or MUAC <125 mm or bilateral pitting edema three months after discharge from SAM or MAM OTP (cohort 3 or Relapse Prevention study).

Secondary study outcomes

For Cohort 1

- The longitudinal prevalence of:
  - MAM defined as the number of MAM diagnoses divided by the total number of monthly visits made by the survey teams.
  - SAM defined by the number of SAM diagnoses divided by the total number of monthly visits made.
- The incidence of:
  - Wasting defined as the number of new cases of wasting diagnosed during the monthly visits made by the survey teams.
  - MAM defined as the number of new MAM cases diagnosed during the monthly visits made by the survey teams.
  - SAM defined as the number of new SAM cases diagnosed during the monthly visits made by the survey teams.
- The prevalence of anemia and severe anemia defined as the proportion of children with a hemoglobin level below 11g/dL and 7 g/dL respectively, at age of 12 months or at the end of the study.
- The mean hemoglobin concentration at age of 12 months or until or at the end of the study.
- The prevalence of stunting defined as the proportion of children with LAZ<-2 (according to the 2006 WHO reference) at age of 12 months or at the end of the study.
- Wasting screening coverage defined as the proportion of children screened (using MUAC, WLZ or bilateral pitting edema) in the month prior to the monthly visit by the interviewers. Two sub-outcomes will also be concerned:
  - Coverage of screening performed by NASGs within one month.
  - Coverage of the family MUAC component, which is the screening performed by caregivers or family member within one month.
- The psycho-cognitive development of the children assessed via the DMC-II score at the end of the study.
- The linear growth rate including:
  - Change in LAZ per month
- Ponderal growth by:
  - The change in WAZ score per month
  - The change in the WLZ score per month
  - MUAC gain (change in MUAC per month)
- Longitudinal prevalence of childhood morbidity, i.e., acute respiratory infections, fever, diarrhea, and malaria, defined as the number of diagnoses of signs of these morbidities divided by the total number of monthly visits made by the survey teams.

- Caregiver knowledge of breastfeeding, complementary feeding, child health and hygiene, screening of wasting using MUAC, MAM and SAM OTP at the end of the study expressed as a cumulative total score and by subdomain.
- IYCF practices:
    - Minimum dietary diversity of children during study follow-up and at the end of the study, defined as the proportion of children who consumed at least 5 of the 8 IYCF food groups (including breast milk) the day before the survey.
    - Minimum meal frequency of children during study follow-up and at the end of the study, defined as the proportion of children who had 3 meals the day before the survey for breastfed children or 4 meals for non-breastfed children. The questions on feeding frequency exclude the consumption of SQ-LNS.
    - Minimum acceptable diet at the end of the study, defined as the proportion of children with both minimal dietary diversity and minimal meal frequency on the day before the survey.
    - Consumption of iron-rich or iron-fortified foods of children at the end of the study. The indicator will be calculated by excluding and including the SQ-LNS as an iron-fortified food.
- The adoption of practices recommended by NASGs and related to WASH, vaccination, and consumption of SQ-LNS at age of 12 months or at the end of the study.

For Cohort 2 [4]

- Nutritional status (measured by WLZ and MUAC) at enrollment in SAM and MAM OTP (pooled), SAM OTP and MAM OTP in children 6-23 months of age.
- Treatment duration of SAM and MAM OTP (pooled), SAM OTP and MAM OTP defined as the number of days spent under treatment (between enrollment and discharge) in children 6-23 months of age in the total sample and in the sample of children who recovered.
- Adherence to the treatment schedule of SAM and MAM OTP (pooled), SAM OTP and MAM OTP in children 6-23 months of age.
- The outcome of SAM and MAM OTP (pooled), SAM OTP and MAM OTP (default, death, transfer to inpatient care) in children 6-23 months of age.
- Average weight gain per study group in grams per day per kg body weight defined by weight change between discharge and admission divided by treatment duration (days) and by initial body weight (kg) for the total sample and the subgroup of children who recovered
- Average MUAC gain per study group in mm per day defined by change in MUAC between discharge and admission divided by treatment duration (days) in the total sample

For Cohort 3

All outcomes will be measured in a sample of children (9-18 month of age) during a home visits three months after successful completing an URENAS or URENAM treatment program

- The prevalence of stunting, defined as the proportion of children with LAZ<-2 (according to the 2006 WHO reference)
- The prevalence of SAM and MAM

[4] All the results cited below come from the registers (including follow-up sheet) of the URENAM/URENAS treatment services (without collecting personal information of mothers/children, such as names, addresses, date of birth, telephone nr.). There will therefore be no data collection from mothers/children attending the consultations.

- Wasting screening coverage defined as the proportion of children screened (using MUAC, WLZ-score or bilateral pitting edema) in the month prior to the interviewer's visit. Two sub-outcomes will also be concerned:
    - Coverage of screening performed by NASGs within one month.
    - Coverage of the MUAC family component, which is the screening performed by a family member within one month.
- The prevalence of anemia and severe anemia defined as the proportion of children with a hemoglobin level below 11g/dL and 7 g/dL respectively, three months after discharge from a SAM or MAM OTP treatment program.
- The average hemoglobin concentration three months after leaving a SAM or MAM OTP treatment program.
- Adoption of IYCF practices (as described for Main cohort) and practices recommended by NASGs related to hygiene (WASH), and consumption of SQ-LNS three months after discharge from a SAM or MAM OTP

Auxiliary results

In addition to these study results, the impact of the intervention on single coverage of SAM and MAM OTP will also be assessed. They are not reported as formal study results because of the imprecision in their measurement.

Two samples of children will be used to assess SAM and MAM OTP single coverage:

a. Main cohort children identified with wasting, SAM and MAM and children recovering from wasting, SAM and MAM
b. Children identified with wasting, SAM and MAM and children recovering from wasting, SAM and MAM during two mass screening campaign in the study villages enrolled in the Main cohort conducted by World Vision Mali towards the end of the study

Result a) will be strongly influenced, in both study groups, by the MAM/SAM cases referred by the study teams who will record anthropometry during monthly visits of the Main cohort. In the case of MAM or SAM, the study teams must, for ethical reasons, refer the children to the health center or CHW for admission to SAM or MAM OTP. Thus, the real impact of, for instance, NASGs (who screen and refer MAM/SAM cases) and the family MUAC approach on enrollment and treatment coverage may be underestimated because study teams identified the wasting episode earlier.

Result b) is obtained from two community-based mass screening campaigns conducted by health staff, CHW and World Vision field workers whose main objective is to screen and refer children suffering from SAM and MAM using the criteria of WLZ, MUAC and bilateral pitting edema. The reason for these campaigns is that previous studies found that using a MUAC only screening approach detects only part of the cases of child wasting. In order not to neglect children with wasting indicated by a WLZ-score<-2, but with a normal MUAC (above 125 mm), these cross-sectional campaigns will screen children in the community using all criteria defining child wasting (MUAC, WLZ, and bilateral pitting edema). These mass screening campaigns will not collect the personal information of children.

Randomization, Sample Size, and Sampling

Randomization

The study area included the health areas covered by 45-46 community health centers in the Koutiala health district in the Sikasso region.

Because of differences between the health areas in terms of population, urbanity/rurality, and remoteness from the nearest health center, the health center catchment areas in the Koutiala district were first stratified by creating mutually exclusive strata using a list of criteria which were:

- Presence of programs related to young children nutrition in the health center catchment area
- Type of health center catchment area (urban/rural)
- Proportion of the population living within 15km of the nearest health center more or less than 10%.

Randomization into comparison and intervention groups was done separately for each stratum. This a priori stratification before randomization ensured a more balanced distribution of health center catchment area characteristics among the study groups.

Subsequently, for the randomization of health center catchment areas, IFPRI's Principal Investigator (LH), (based in the United States) first generated a list of random and unique two-letter codes (Stata 16.0 software) that were allocated to the study arms (23 for control and 23 for intervention) balancing by stratum. The two-letter code lists (without study arm allocation information) were sent to Mali, where, per stratum, the program implementer (World Vision Mali) prepared a corresponding number of opaque envelopes each containing one of the two-letter codes from this list.

Before the drawing ceremony, World Vision Mali also prepared 46 pieces of paper of identical shape containing each a name of the 46 health center catchment areas. These papers were folded in an identical manner to hide the names of the health center catchment areas and for each stratum placed in 4 bags (one bag per sampling stratum).

A draw ceremony was organized in the office of World Vision in Koutiala. The ceremony was attended by the Chief Medical Officer of the health district (secretary), the World Vision Mali coordinator, two community representatives acting as witnesses of the process, the president of the FELASCOM (District-level association of health center management committees) in charge of the draw, and the IFPRI principal investigator (by videoconference) who was the only one to have the list of two-letter codes and the assignment to the study arms.

For each stratum, the president of FELASCOM drew a piece of paper (containing the name of a health center catchment area) from the bag and drew also one of the envelopes (containing the unique two-letter code). As he unfolded the piece of paper, he read aloud the name of the health center catchment area and showed the name of written on the piece of paper to all participants and wrote that name on the front of the envelope. This process continued until all 46 envelopes were linked to a health center catchment area. After this step, the IFPRI principal investigator projected the two-letter codes with study arm assignment to all participants. During this last step, the envelopes were opened one by one and the president of the FELASCOM read the name of the health center catchment area and the two-letter code aloud so that the result could be written by the secretary on a flip chart and made visible to all participants.

Sample Size

Sample sizes were calculated prior to randomization and data collection. However, as mentioned earlier, 3 study clusters were omitted from the study during the baseline survey because of severe security threats. In addition, the sponsor allowed the study to be extended by two months to the end of February 2022. We therefore recalculated the sample size for each of the primary outcomes for 42 study clusters instead of the original 45 and including two additional months of recruitment of study subjects.

Cohort 1 - Main Outcome: Longitudinal prevalence of wasting

Scenario with 45 clusters available (SAP version 1.0)

Considering a baseline longitudinal prevalence of 13.2% (obtained from the control group of the PROMIS Mali study) with a standard deviation (SD) of 0.239 for the control group and an SD of 0.189 for the treatment group, a mean cluster size of 35 (anticipating a loss of follow-up of 15%), a statistical power of 80%, a type I error of 5% and an intra-cluster coefficient of 0.04, we need at least 44 clusters to detect a minimum difference of ~ 4.84 percentage points in the 6-14 months age group.

Scenario with 42 clusters available (SAP version 2.0)

Considering a baseline longitudinal prevalence of 13.2% (obtained from the control group of the PROMIS Mali study) with a standard deviation (SD) of 0.239 for the control group and an SD of 0.189 for the treatment group, a mean cluster size of 55 (anticipating a loss of follow-up of 15%), a statistical power of 80%, a type I error of 5% and an intra-cluster coefficient of 0.04, we need at least 42 clusters to detect a minimum difference of ~ 4.55 percentage points, , which is equivalent to a relative risk of 0.65.

Cohort 2 - Primary Outcome: recovery Rate

The aim is to include all URENAS and URENAM treatment data for children between 6 and 23 months of age. Using 2019 health district statistics, the total number of SAM and MAM cases (6-23 months of age) over 8 months of follow-up was estimated at 750 and 1,000 respectively.

Scenario with 45 clusters available (SAP version 1.0)

With 44 clusters that each have an average of 35 cases of SAM and MAM over a 7-month period, assuming a URENAS/URENAM cure rate of 79%, a statistical power of 80%, a type I error of 5% and an intra-cluster coefficient of 0.05, a minimum difference of ~9.0 percentage points in the URENAS/URENAM cure rate could be detected in children aged 6-23 months.

Scenario with 42 clusters available (SAP version 2.0)

With 42 clusters that each have an average of 50 cases of SAM and MAM over a 7-month period, assuming a URENAS/URENAM cure rate of 79%, a statistical power of 80%, a type I error of 5% and an intra-cluster coefficient of 0.05, a minimum difference of ~9.0 percentage points in the URENAS/URENAM cure rate could be detected in children aged 6-23 months.

Cohort 3 - Primary Outcome: Prevalence of Relapse

Scenario with 45 clusters available (SAP version 1.0)

We considered a 35% prevalence of relapse in children aged 9-17 months, three months after treatment completion (estimate obtained in the control group of the PROMIS Mali study). We need a minimum of 44 clusters with a mean cluster size of 20 (anticipating a 15% loss of follow-up), a statistical power of 80%, a type I error of 5% and an intra-cluster coefficient of 0.04 to detect a minimum difference of ~ 12.0 percentage points in the 9-18[5] months age group.

Scenario with 42 clusters available (SAP version 2.0)

We considered a 35% prevalence of relapse in children aged 9-17 months, three months after treatment completion (estimate obtained in the control group of the PROMIS Mali study). We need a minimum of 42 clusters with a mean cluster size of 20 (anticipating a 15% loss of follow-up), a statistical power of 80%, a type I error of 5% and an intra-cluster coefficient of 0.04 to detect a minimum difference of ~ 13.5 percentage points in the 9-18 months age group.

Sampling and study inclusion criteria

Cohort 1

In May 2021 a census will be organized to enlist all households with at least one child aged 0-6 months in 3-6 villages per health center catchment area. These villages will be selected proportion to their population size. Between May and November 2021, a random sample of children meeting the study's inclusion criteria will be enrolled using these census lists.

The criteria for inclusion of children in Cohort 1 are:

- 6-6.9 months of age;
- Singleton child
- The mother does not intend to leave the study area before March 2022;
- The child does not have congenital malformations that make anthropometric measurements impossible.

Cohort 2

Anthropometry and participation data for the MAM and SAM OTP of all malnourished children between 6 and 23 months of age will be collected from the MAM and SAM OTP anonymously (i.e. personally identifiable information such as name, address and date of birth will not be collected) for mothers-children from villages not selected for cohort 1. For mothers-children from villages selected for cohort 1, informed consent to participate in the study of cohort 2 and potentially cohort 3 (if randomly selected for cohort 3) will be obtained before recording the personally identifiable information.

Cohort 3

[5] This age range corresponds to children discharged from URENAM/URENAS treatment between 6 and 15 months (the follow-up measure for cohort 3 will be done 3 months after discharge).

A random sample of children aged 6-15 months will be selected from the lists of children declared recovered from wasting from URENAM and URENAS treatment between May and December 2021. These children will be home visited by a fieldworker 3 months after being declared cured and discharged from treatment to diagnose a possible relapse.

The criteria for inclusion of children in Cohort 3 are:

- Child ex-MAM or ex-SAM at least 3 months after being declared recovered from wasting through URENAM/URENAS treatment
- Singleton child
- Child aged 6-15 months when successfully completing of URENAM / URENAS treatment.
- The mother does not intend to leave the study area within 4 months after the discharge of her child from the URENAM / URENAS treatment
- The child does not have congenital malformations that make anthropometric measurements impossible.

Measurements and indicator compilation

For this study, interviewers use Android tablets with specialized software for computer interviews (Survey Solutions software).

In the Main cohort and the Relapse prevention study, children's anthropometric measures are recorded during home visits, and several modules of questions were administered to provide essential insight into relevant parental and household practices. The preferred respondent is the main caregiver of the target child. All interviews are conducted in the respondent's language.

For the Treatment cohort, each nutritional unit in health centers (42) is visited by an enumerator once a week and each community treatment unit (CHW) is visited monthly. The enumerator is responsible for monitoring data quality and copying data from the registers to electronic forms on their tablet.

Questionnaires

- Village questionnaire: A module of questions on infrastructure, health services present in the village, external shocks and the security situation was administered to the village representative (or another knowledgeable person).
- Household level: with questions about household composition, educational level, socio-economic status, food security, livestock/land/asset holdings and household construction materials. There was also a module of questions on household WASH practices and household food security
- Caregiver and child level: There were questions on parenting style, primary occupation, postpartum depression, and nutritional status of the caregiver, as well as questions on child health and nutrition of the child. Specific questions on IYCF and WASH knowledge and practices will be asked to the caregiver in charge of the child's care. Other modules will include the use of (community-based) health care services and number of contacts with the NASGs. Questions on fine motor, gross motor and communication and language development of the young child will be asked

to the caregiver by study enumerators who will also directly observe the child's skills using the DMC-II tool[6].

Date of birth to calculate the child's age at study enrolment will be obtained from, in order of priority: birth certificate, vaccination card or by using a local events calendar. Child anthropometry is taken at enrolment and at each monthly visit of the longitudinal follow-up in the Main cohort and three months after discharge from SAM and MAM OTP for the Relapse prevention study. The child's weight is taken using an electronic scale (SECA 876, Germany) to the nearest 100g. The length is measured to the nearest 1 mm using length boards. The Mid-Upper Arm Circumference (MUAC) was measured using a non-stretch tape with 0.1 cm accuracy (SECA 201, Germany). All measurements are taken in duplicate by a trained anthropometrist and the study enumerator as assistant. All measurements are standardized prior to the study. Standardization exercises using repeated measurements on 10 children aged 6-23 months will be repeated every two months during the study. The MUAC, weight and height of the mother are also measured using the same equipment. The child weight-for-height and height-for-age z-scores will be calculated using the *zscore06* command in Stata[7], which uses the World Health Organization's growth standard[8]. A wasting episode is defined as follows: it starts when a child is found to be wasted at the monthly measurement by study teams and ends when a child does not suffer from wasting at one monthly measurement. MAM and SAM episodes are defined the same way.

Child hemoglobin concentration is measured at enrolment and end of study follow-up in Main cohort, as well as in children enrolled in the Relapse prevention study. For this purpose, finger blood is collected. The hemoglobin concentration in the second drop of finger blood is measured by spectrophotometry using a HemoCue device 301 (HemoCue Ltd, Dronfield, UK). Anemia in children and severe anemia are defined by a hemoglobin concentration of less than 11 g.dl-1 and less than 7 g.dl-1 respectively.

During the monthly home visits study enumerators ask the caregiver to recall any child morbidity symptoms (acute respiratory infections, cough, vomiting, breathing difficulties, diarrhea, and fever) over the 3 days preceding the monthly visit. A diarrheal episode is defined as at least three loose stools in the last 24 hours or stools with blood. Fever is measured with a standard thermometer. The presence of malaria parasites is assessed using CareStart™ Malaria Pf/Pv Combo rapid diagnostic test in capillary finger blood if the child's body temperature is above 37.5°C or if the mother reports a fever episode in the child within the past 48 hours. Morbidity outcomes are be expressed as a longitudinal prevalence by taking the proportion of days with symptoms during the recalled period in days.

Data on service coverage by NASG and the local health system, as well as coverage of screening of wasting are collected through caregiver recall.

[6] Prado, Elizabeth L. et al. 2014. "Extending the Developmental Milestones Checklist for Use in a Different Context in Sub-Saharan Africa." *Acta Paediatrica, International Journal of Paediatrics* 103(4): 447–54

[7] Leroy, J.L. 2011. "ZSCORE06: Stata Command for the Calculation of Anthropometric z-Scores Using the 2006 WHO Child Growth Standards." https://econpapers.repec.org/software/bocbocode/s457279.htm (September 30, 2017)

[8] World Health Organisation. 2006. "WHO Child Growth Standards." *Acta Paediatr Suppl* 95(Supplement 450): 1–104

Wasting, MAM and SAM OTP single coverage are estimated using the observed number of uncovered wasting, MAM and SAM cases, observed number of covered recovering cases and an estimate of uncovered recovering cases[9]:

$$single\ coverage = \frac{C_{in} + R_{in}}{C_{in} + R_{in} + C_{out} + R_{out}}$$

where C_{in} = current wasting, SAM or MAM cases in SAM or MAM OTP, C_{out} = current wasting, SAM or MAM cases not in SAM or MAM OTP, R_{in}= recovering wasting, SAM or MAM cases in SAM or MAM OTP, and Rout = recovering wasting, SAM or MAM cases not SAM or MAM OTP estimated for World Vision screening campaigns as $\frac{1}{k} \times (R_{in} \times \frac{C_{in}+C_{out}}{C_{in}} - R_{in})$ and k = the mean length of a recovered episode of wasting, assumed to be 5.1 months for MAM and 5.7 months for SAM based on a previous longitudinal study conducted in an adjacent region[10]. For single coverage estimate using the Main cohort data, R_{out} is estimated by observing recovery between two monthly measurements.

Data management

The management of all research data will follow IFPRI's institutional research data management protocol, overseen by the IFPRI Data Governance Committee, as well as UNICEF data governance standards. All received data will be carefully anonymized and de-identified so that the privacy of participants and research subjects is fully protected. Personally Identifiable Information data (PII; names, telephone numbers, GPS coordinates, date of birth) will be stripped from the research data when receiving the data on the server. During data collection, PII data will be stored in a separate secured server folder accessible only to the Principal Investigators. PII data will only be used to identify a household to conduct additional verification of previously collected data. When data collection and data validation have been completed, PII data will be destroyed.

In accordance with IFPRI's policy on research data management and open access, at the time of publication of scientific articles presenting the primary results, fully anonymized databases will become a public good and will therefore be made available to the scientific community, government, partners.

All data collection is scheduled to start once the study trial is registered on clinicaltrials.gov.

Statistical analysis

Data sources

The training of NASG groups is scheduled to be concluded in the last week of July 2021. Therefore, cohort 1 and 2 data from May to July 2021 will be used to assess the study baseline and to inspect imbalances in sample characteristics between study arms. Follow-up data from August 2021 to February 2022 will be

[9] Balegamire S, Siling K, Alvarez JL, Guevarra E, Woodhead S, Norris A. 2015. "A Single Coverage Estimator for Use in SQUEAC, SLEAC, and Other CMAM Coverage Assessments." *Field Exchange, vol. 49*

[10] Barba, Francisco M., Lieven Huybregts, and Jef L. Leroy. 2020. "Incidence Correction Factors for Moderate and Severe Acute Child Malnutrition From 2 Longitudinal Cohorts in Mali and Burkina Faso." *American Journal of Epidemiology* 189(12): 1623–27

used to assess the impact of the program. For cohort 3, data from October 2021 to February 2022 will be used to study the impact of the intervention on relapse.

Guiding principles and analysis strategy

Data will be analyzed on intent-to-treat basis. To allow for an analysis "as randomized" in the presence of missing data, we will impute missing data of the primary outcomes of the longitudinal study using a multiple imputation strategy under the *missing at random* assumption. Data management, data cleaning, and statistical analyses will be done using Stata 17.0 (Statacorp, USA). The statistical significance will be set at 5%. All statistical tests will be two-sided. The analysis of primary study outcomes will not be adjusted for multiple testing (n=3 primary study outcomes) given the data involved represent three distinct experimental and control study arms with each one primary study outcome.

We present results from non-adjusted and adjusted linear mixed effects regression models. The non-adjusted models use *health center catchment area* as a random intercept and sampling stratum as a fixed effect to represent the study design. Adjusted regression models further include as fixed effects the cluster baseline means (May-July 2021 period) of the outcome under analysis, month of enrollment into the study (cohort effect), child sex, child age, and primiparity. Furthermore, if any baseline household, caregiver of child characteristic is imbalanced (absolute difference of ~5pp) between study groups, we conduct an additional robustness analysis adjusting the analysis for this covariate.

Cohort 1 specific considerations

Children are enrolled in the study cohort at the age of 6 months, when the program foresees MUAC screening measurements and SQ-LNS distribution by NASG to start. From August 1th, 2021 (start of the program) onwards, NASG in both intervention and comparison study arms are tasked to conduct home visits to pregnant and lactating women and caregivers of children up to 24 months of age to provide BCC on child nutrition, health and WASH. Therefore, it is likely that caregivers with children of 6 months enrolled in the cohort after the start of the program have been exposed to BCC prior to study inclusion. This exposure to BCC may have been greater in the intervention group because in that study group the program formed new NASGs proportionate to village population size (contrary to the comparison group with typically counts 1-2 NASG per village independent of population size). For this reason, all analyses related to study outcomes on the impact pathway related to BCC (knowledge indicators, IYCF, immunization rates, WASH practices, child anthropometry) include the data at enrollment at the age of 6 months. For study outcomes related to activities that only start between the age of 6 and 7 months, like MUAC-based screening and the distribution of SQ-LNS, we consider data related to the program's impact from the child age of 7 months onwards. To assess the intervention's impact between study enrollment (child age of 6 months) until the end of the study (9 to 12 months of age)

Treatment of missing data for primary study outcomes

We conducted multiple imputation of missing longitudinal anthropometric data using fully conditional specification, also known as multiple imputation by chained equations which imputes missing values

under the missing at random assumption[11]. Imputation variables were the wide-form anthropometric measurements and child age, and complete predictors (independent variables) were child sex and randomization stratum. We used 50 imputations with 10 burn-in iterations to allow for convergence to a stationary distribution before the imputation. Imputation quality will be assessed using trace plots by iteration to assess convergence and using box plots to compare distributions of observed and imputed values.

Statistical analysis methods

For the analysis of the longitudinal prevalence of wasting (primary outcome), severe and moderate acute malnutrition and morbidity, the incidence of acute malnutrition and wasting we use mixed-effects Poisson regression models with robust estimation of standard errors to provide adjusted risk estimates. Results from non-adjusted and adjusted regression models will be presented as described under the paragraph on 'Guiding principles and analysis strategy'.

Continuous outcomes (such as WLZ, MUAC, LAZ) that are measured monthly will be regressed over calendar date using a linear mixed-effects model including *health center catchment area* and *child* as a random intercept and *follow-up time* as a random slope. For each model, we assess if the addition of a quadratic term of time improves the model fit significantly by applying a likelihood ratio test comparing the model with and without the quadratic term. To assess the specific impact of the intervention (SQ-LNS, MUAC screening and BCC) from enrollment at the age of 6 months to the end of follow-up, we assess the interaction between intervention allocation and time.

Dichotomous outcomes will be analyzed using linear probability mixed-effects models with robust estimation of the standard errors with *health center catchment area* as random effect and sampling stratum as fixed effect to demonstrate the impact of the intervention on the dichotomous outcome averaged over the follow-up period or at the end of the study. For repeated measures, to assess the specific impact of the intervention (SQ-LNS, MUAC screening and BCC) from enrollment at the age of 6 months to the end of follow-up (cohort 1), we assess the interaction between intervention allocation and follow-up time using a linear probability mixed-effects model with *health center catchment area* and *child* as a random intercepts and *follow-up time* as random slope.

Considerations for analysis

Baseline characteristics

For all variables measured, the available values at enrollment and collected from May to July 2021, i.e. before the start of the intervention, will be considered as baseline characteristics. The background characteristics of subjects who completed the study will be presented by intervention group. Following CONSORT guidelines, no statistical tests will be conducted comparing baseline characteristics between

[11] De Silva AP, Moreno-Betancur M, De Livera AM, Lee KJ, Simpson JA. (2017) A comparison of multiple imputation methods for handling missing values in longitudinal data in the presence of a time-varying covariate with a non-linear association with time: A simulation study. BMC Med Res Methodol; 17:1–11

study arms. Differences in baseline characteristics will be appreciated by comparing values of means and proportions.

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 are available.
- Continuous variables: mean and SD or median and interquartile range, as appropriate.

Characteristics	Cohort	Comments/References
Household level		
HH size (mean/SD)	1,3	
Number of under 5 children in household (mean/SD)	1,3	
Number of adults (15-64 y) (mean/SD)	1,3	
Polygamous HH(%)	1,3	
HH food insecurity (mean and proportion of food secure HH)	1,3	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
SES status classification (tertiles, %)	1,3	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, and home ownership. Tertiles of the first principal component (i.e., the one with the highest eigenvalue) will be used.
Improved water treatment technologies used (%)	1,3	USAID (2021) Water and development- Indicator handbook. Washington DC
Improved sanitation facility (%)	1,3	USAID (2021) Water and development- Indicator handbook. Washington DC
Handwashing station with soap available (%)	1,3	USAID (2021) Water and development- Indicator handbook. Washington DC
Improved primary water source (%)	1,3	USAID (2021) Water and development- Indicator handbook. Washington DC
Distance to nearest health center, km	1,3	Linear distance between households and the nearest health center calculated using Global Position System coordinates
Distance to nearest Community health worker, km	1,3	Linear distance between households and the nearest Community health worker calculated using Global Position System coordinates. Limited to villages that are covered by a community health worker.
Household head level		
Male head of household (%)	1,3	
Adult (15-64 years) (%)		
Elderly (65+ years) (%)		
Paternal level		
Father is head of HH (%)		
Age (mean/SD)	1,3	
Male (%)	1,3	

Head of household school attendance	1,3	
Has an income generating activity (%)	1,3	
Parental confidence scale	1,3	Črnčec R, Barnett B, Matthey S. Development of an instrument to assess perceived self-efficacy in the parents of infants. Res Nurs Health. Res Nurs Health; 2008
Severe clinical range (less than 31)		
Moderate clinical range (31-35)		
Maternal		
Biological mother (%)		
Age (mean/SD)	1,3	
Spouse of head of household (%)	1,3	
School attendance (Never attended school, %)	1,3	
Has an income generating activity (%)		
Height (meand/SD)	1,3	
Possible depression (EPDS≥10)	1,3	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.
Parental confidence scale	1,3	Črnčec R, Barnett B, Matthey S. Development of an instrument to assess perceived self-efficacy in the parents of infants. Res Nurs Health. Res Nurs Health; 2008; 31:442–53.
Severe clinical range (less than 31)		
Moderate clinical range (31-35)		
Number of food groups consumed over last 24h (mean/SD)	1,3	Out of a maximum of 10 food groups as proposed by Martin-Prevel Y, Allemand P, Wiesmann D, Arimond M, Ballard T, Deitchler M, et al. Moving forward on choosing a standard operational indicator of women's dietary diversity. Rome: FAO; 2015. Food groups are starchy staple foods, nuts and seeds, flesh foods, dark green leafy vegetables, pulses, dairy, eggs, vitamin A rich fruits and vegetables, other vegetables and other fruits
Minimum dietary diversity (having consumed at least 5 of 10 food groups during the last 24h)	1,3	Martin-Prevel Y, Allemand P, Wiesmann D, Arimond M, Ballard T, Deitchler M, et al. Moving forward on choosing a standard operational indicator of women's dietary diversity. Rome: FAO; 2015
Body Mass index (kg.m-2) (mean/SD)	1,3	
Underweight (BMI<18.5) (mean/SD)	1,3	
Assisted delivery of last pregnancy (%)	1,3	
Number of antenatal consultations attended during last pregnancy (%)	1,3	
Breastfeeding, complementary feeding, health and hygiene, MUAC screening, CMAM knowledge score (mean/SD)	1,3	Self-developed questionnaire based on training curricula for GSAN. Mean scores are presented by domain, as well as an overall score based on the 36 questions
Child		
Age (mean/SD)	1,2,3	
Male (%)	1,2,3	

Primiparity	1,3	
Minimum dietary diversity (%)	1,3	Indicators for assessing infant and young child feeding practices: definitions and measurement methods. Geneva: World Health Organization and the United Nations Children's Fund (UNICEF), 2021
Minimum meal frequency(%)	1,3	
Minimum acceptable diet(%)	1,3	
Consumption of iron-rich or iron-fortified foods (%)	1,3	
Hemoglobin concentration (%)	1,3	As measured by hemocue 301
Anemia (hb<7 g/dl)	1,31,3	
Severe anemia (hb<11 g/dl)		
Weight-for-length Z-score (mean/SD)	1,2,3	Relative to WHO 2006 growth reference
Wasting (WLZ<-2 or edema)	1,2,3	Relative to WHO 2006 growth reference
Severe Wasting (WLZ<-3 or edema)	1,2,3	Relative to WHO 2006 growth reference
MUAC (mean/SD)	1,2,3	Relative to WHO 2006 growth reference
Acute malnutrition (WLZ<-2 or MUAC<125mm or edema)	1,2,3	Relative to WHO 2006 growth reference
Severe Acute malnutrition (WLZ<-5 or MUAC<115mm or edema)	1,2,3	Relative to WHO 2006 growth reference
Length-for-age Z-score (mean/SD	1,2,3	Relative to WHO 2006 growth reference
Stunting (LAZ<-2)	1,2,3	Relative to WHO 2006 growth reference
Weight-for-age Z-score (mean/SD	1,2,3	Relative to WHO 2006 growth reference
Underweight (WAZ<-2)	1,2,3	Relative to WHO 2006 growth reference
Child development score (DMC-III)	1	Prado, E.L. , Abubakar, A.A. , Abbeddou, S. , Jimenez, E.Y. , Some, J.W. , & Ouedraogo, J.B. (2014). Extending the developmental milestones checklist for use in a different context in Sub-Saharan Africa. Acta Paediatrica, 103,
Gross motor (mean Z-score/SD)	1	Z-scores(individual value minus mean group value divided by the SD) will be calculated using data during the baseline period of the study
Fine motor (mean Z-score/SD)	1	Z-scores(individual value minus mean group value divided by the SD) will be calculated using data during the baseline period of the study
Language (mean Z-score/SD)	1	Z-scores(individual value minus mean group value divided by the SD) will be calculated using data during the baseline period of the study

Effect modification

Studying effect modification of an intervention offers important insights in the mechanism of a possible impact. In this study, we applied a stratified cluster randomized sampling, stratifying by health center catchment area characteristics (urban/rural, proportion of population living >15km from a health center, previous activities on child nutrition delivered by World Vision Mali). To gain more operational insight on how the intervention performed according to stratum characteristics we will assess possible effect modification of the intervention by these strata. Given that these tests are explorative in nature, we assess

interaction terms at a statistical significance set at 10%. Statistical models are the same as described in earlier sections. Results from both adjusted and non-adjusted regression models will be presented. If there are indications of effect modification by stratum characteristics, we will present impact results stratified by stratum.

Sensitivity analyses

We will assess whether impacts on primary outcomes in cohort 1 and 3 are sensitive to the definition of wasting, by assessing impact on wasting prevalence and incidence when wasting is defined with 2 alternative definitions: 1) MUAC < 125 mm or oedema, and 2) WLZ < -2 Z-scores or oedema.