Study protocol

16 August 2021

Table of Contents

The uploaded data is primary data that was collected as part of a long-term outcome study conducted in Sokoto state, Nigeria in 2018 and 2019. This protocol describes the objectives and methodology of the research, as well as the uploaded datasets.

Title	Incidence of severe acute malnutrition after treatment: A prospective matched cohort study in Sokoto, Nigeria
Study type	Prospective matched cohort study
Data producer and depositor	Oxford Policy Management Ltd.
Data production date	22 July 2021
Citation requirement	Oxford Policy Management. Incidence of severe acute malnutrition after treatment: A prospective matched cohort study in Sokoto, Nigeria, Version 1.1 of the public use dataset (November 2021). Downloaded from [url] on [date].
Funding source	The study was funded by the Children's Investment Fund Foundation. Grant/Award Number: G-1603-01140.
Disclaimer	The user of the data acknowledges that the original collector of the data, the authorised distributor of the data, and the relevant funding agency bear no responsibility for use of the data or for interpretations or inferences based upon such uses.
Data confidentiality	The uploaded data have been anonymised and are accessible for statistical and research purposes only. The data will be used solely for reporting of aggregated information, and not for investigation of specific individuals or facilities. No attempt will be made to re-identify respondents, and no use will be made of the identity of any person or establishment discovered inadvertently. Any such discovery would immediately be reported to Oxford Policy Management Ltd.
Related publication	The uploaded data is associated with a journal article. Citation of the journal article: Adegoke, O, Arif, S, Bahwere, P, et al. Incidence of severe acute malnutrition after treatment: A prospective matched cohort study in Sokoto, Nigeria. *Matern Child Nutr*. 2021; 17:e13070. https://doi.org/10.1111/mcn.13070.

Background

Acute malnutrition is an important concern for children aged under 5 years, both globally and in Nigeria. Since the 2000s and the advent of Ready-to-Use Therapeutic Foods (RUTF), Severe Acute Malnutrition (SAM) has been commonly addressed through the WHO-recommended approach of *Community-based Management of Acute Malnutrition* (CMAM). CMAM programs provide treatment for children aged 6-59 months through an outpatient therapeutic program (OTP) service for uncomplicated cases and an inpatient service for complicated cases. Volunteers are tasked with actively finding and referring wasting cases in their communities and with following-up children when they either drop out or complete the program, to minimize the likelihood of relapse. In some contexts, treatment and support for moderate acute malnutrition (MAM) is also provided.

In Nigeria, children are admitted and discharged from the CMAM program using mid-upper arm circumference (MUAC) as the main criterion. Children with a MUAC <115 mm and no apparent health complications are admitted into the OTP.

Evidence exists that children who have gone through an episode of SAM remain at higher risk of morbidity and mortality than children who have not. However, evidence on SAM relapse is sparse with a high variation in estimates in the literature due to both contextual and methodological differences. This makes it difficult to understand the persistent risk of a SAM episode after initial recovery from the CMAM program, as well as the associated risk factors. Another important gap in the literature is the absence of comparison groups, making it difficult to determine the excess risk for SAM associated with a recent SAM episode.

Programmes working to treat SAM, such as the CMAM approach, should focus both on short-term survival and on improving long-term outcomes.

Research objectives

The primary objective of this study was to assess the persistent and excess risk of SAM among children treated by the CMAM program.

- The persistent risk was assessed by measuring the 6-month incidence rate of relapse into SAM among children discharged as cured from the OTP services of the CMAM program.

- Excess risk was assessed by comparing this rate of relapse to the 6-month incidence rate of SAM in a cohort of community controls.

Our secondary objective was to identify factors that are associated with the risk of relapse.

Study design

This prospective matched cohort study was conducted from September 2018 to May 2019 in five rural local government areas (LGAs) in Sokoto State, northern Nigeria. Nine out of the 23 LGAs in Sokoto State hosted the CMAM program. Of these, four were excluded because they were either hosting another study looking at improving CMAM delivery, were peri-urban, or not easily accessible. In the selected five LGAs, the CMAM program had been running since 2010 and was being implemented throughout the study period. Within each LGA, five health facilities were hosting the CMAM program, and the study covered this exhaustive list of 25 health facilities (i.e. there was no sampling of health facilities within LGAs, all health facilities that were hosting the CMAM program in each LGA were included in the study).

The study followed two cohorts of subjects:

1. First, a cohort of children who were i) admitted into OTP and discharged alive and as **cured**[1], ii) were aged 6-59 months at admission into the OTP, iii) had not previously been admitted to the OTP or inpatient care, and iv) whose households resided in the

[1] Children are discharged as cured from OTP if their MUAC is superior or equal to 115mm, and there are no signs of bilateral pitting oedema.

catchment area of the selected health facilities[2] and were not planning to move out. ➔ These children are subsequently referred to as 'OTP-cured children.'

2. Second, a cohort of children i) from the same communities as OTP-cured children who had ii) no history of SAM or treatment for SAM, iii) no anthropometric or clinical signs of MAM (MUAC <125 mm) or of severe stunting (height/length-for-age z-score (HAZ/LAZ) <-3 SD) at the time of recruitment into the study, and iv) who were matched to the OTP-cured children based on a set of criteria. Each community control child was matched to an OTP-cured child based on the following criteria: i) residence (living in the same community), ii) age in months (being of similar age, and allowing up to 3-month difference), iii) sex (having same sex), iv) age of the mother (below or above 20 years of age), and v) level of education of the mother (no education, completed primary, and completed secondary or above). ➔ These children are subsequently referred to as 'community control children.'

Exclusion criteria for both cohorts of children for enrolment into the study included:

- Presence of disability or any congenital disease (after clinical examination) that affects growth or prevents accurate anthropometric measurement and/or prevents the child from eating normally;
- A sibling already enrolled into the study;
- The biological mother of the child having passed away; or
- Having a mother <15 years old.

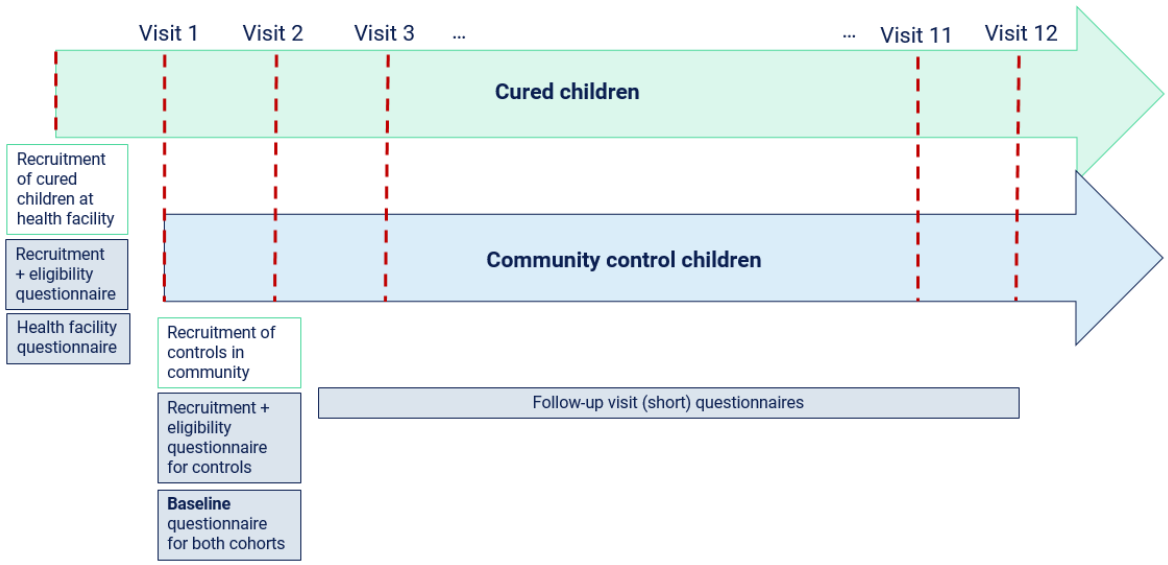
Cohort timeline

This cohort study was implemented in several phases that took place sequentially, and that included recruitment of OTP-cured children, recruitment of community control children, first home visit (to collect baseline characteristics), and subsequent follow-up home visits. At each phase, different questionnaires were administered.

At each health facility, a CMAM day is held once a week, rotating through all facilities per LGA so that no two facilities within an LGA have a CMAM day on the same day per week. Children who are enrolled in the CMAM programme and their caregivers attend CMAM days for check-ups and treatment. This study recruited OTP-cured children at these CMAM days, identifying children who were discharged as cured.

[2] The data collection teams worked with the health facility and CMAM staff to make a list as exhaustive as possible of the communities in the catchment area of the health facility. While the CMAM programme in practice does not exclude children based on their community (for example, many children from Niger travel far with their mothers to benefit from the CMAM programme), the teams only had the authorisation and resources to follow up with children within the catchment area of the study health facilities.

Figure 1: Cohort timeline



The data collection timeline was as follows (Figure 1 provides a graphical summary):

1. OTP-cured children were recruited at health facilities on a rolling basis between September and November 2018. There, the study team screened **all** children who had been discharged as cured on that day for eligibility and consent to participate in the study. This meant enrolling children as they were successfully discharged from the CMAM programme in the 25 health facilities that formed part of this study. At each CMAM day and health facility, teams of two interviewers were present throughout the day to ensure that all children discharged from the programme on that day were screened for possible enrolment into the study and recruited if eligible up until the minimum sample size is reached.

2. Following recruitment into the study, each OTP-cured child was tracked to their community and visited at their home within 2 to 3 weeks of their initial recruitment. Field teams used the information collected at recruitment to locate children in their community. Most communities were accessed either on foot or by motorcycle. This constituted the first home visit where a long **baseline** questionnaire was administered.

3. Immediately after the first home visit of each OTP-cured child, a search for a suitable community control for that child was conducted. For each OTP-cured child, potential community controls were identified using a snowball approach. In essence, this approach meant that interviewers were referred to potential community control children by the mother of the OTP-cured child. Potential community controls were assessed with respect to their eligibility to enter the study and to whether they matched the corresponding OTP-cured child on the set of criteria mentioned earlier. The first community control to meet both sets of criteria was recruited into the study and the search for a community control for a given OTP-cured child ended at that stage. Once a control child was identified, the same first home (baseline) questionnaire was administered to the household and mother of that child.

4. Afterwards, both cohorts were followed-up fortnightly for a total of 12 home visits (the 12 visits includes the first baseline home visit).

5. Participation in the study for both cohorts ended at the 12th home visit, unless a child developed SAM earlier or dropped out of the study (e.g. family no longer consented to participate in the study or moved out of the community, or child had died), whichever came first. In total, children were followed up for a duration of up to six months after discharge from OTP. **Within the six months of follow-up, if children were identified as being SAM by the field team, they proceeded to exit the study** and interviewers referred those children to CMAM services.

Questionnaires and data collection tools

During recruitment of OTP-cured children at the health facility

- A **recruitment questionnaire for OTP-cured children** was administered to the mother of the child on the day children were discharged and recruited into the study. This questionnaire assessed eligibility of the child and collected some information to help with locating the home of the child.

- Additionally, data on children's health status at admission and discharge from the OTP were also collected from registration and treatment tracking cards kept at the facility by staff (**OTP cards and Ration cards**). This data was scanned on enumerator's tablets and later on entered into a database by data entry staff. Information that was entered from these records included anthropometric measurements and morbidity at admission, duration of treatment and anthropometric measurements at each visit to the OTP. Note that data from the scanned OTP and Ration cards has not been uploaded for public use due to data quality concerns. The data suffers from many missing observations, given that this data was not directly collected by the enumerators but relied on health facility staff filling in the OTP and Ration cards for the treated children.

- During this phase, a **health facility questionnaire** was also administered in each health facility to assess adherence of the health facility to the Nigeria CMAM national guidelines and availability of OTP-related drugs and equipment and the general quality of infrastructure and resources. The survey also collected data on shocks that affected the catchment area of facilities in the year prior to the survey, such as drought, floods, sandstorms, and security-related events. In each health facility, this survey was administered once, on the first day the interview team visited the health facility. The survey used direct observation as well as interviews with the head of the health facility and the CMAM focal person in charge. If either of these individuals were not available on the day, other knowledgeable health facility member was asked to respond to the questions.

During recruitment of community control children

- A **recruitment questionnaire for community control children** was administered to the mothers of the children to assess eligibility and matching criteria and decide if they can be recruited.

During the first home visit

- At the first home visit, a long **baseline questionnaire** was administered to the mother of the recruited child and the household head to collect baseline information across several domains related to the child, mother, and household (Table 1). Children's MUAC was measured using the WHO/UNICEF-recommended MUAC tape and measurement protocol, whereas height and length were measured with a precision of 0.1 cm, using boards manufactured by SECA: standing boards for children who were able to stand and lying-down boards for children unable to.

Table 1: Domains included in first home visit questionnaire

Level	Dimension
Child level	Height/length, mid-upper arm circumference
	Demographics
	Breastfeeding history
	Co-morbidities in the 2 weeks prior to the survey
	Immunization status
	Dietary diversity (24 hours prior to survey)
Mother level	Demographics
	Economic activity and education status
	Knowledge on child feeding and health-seeking behaviour
	Reproductive history and care
	Perceived OTP experience
	Networks in community
Household level	Household demographics and composition
	Economic activity and education of household head
	Household assets and wealth
	Water, sanitation, and hygiene infrastructure
	Household food security and dietary diversity
	Deaths in the household in the year prior to the survey

<u>During the follow-up home visits</u>

- At each follow-up home visit, a short **follow-up questionnaire** was administered to the mother of the child to collect child-level co-morbidity data in the 2 weeks preceding the visit (a subset of the questions asked in the baseline questionnaire) and to measure the child's MUAC.

- In the final follow-up visit (i.e. the visit when the child exited the study either because they developed SAM or if they reached the final 12th visit), additional questions were asked of the household including on mother's employment status, changes in the breastfeeding and pregnancy status of mothers, deaths in the household, household food security, child feeding, household and child dietary diversity, and mother's feeding knowledge and practices. These questions were a subset of those asked in the baseline questionnaire and they were added in order to understand if household, mother, or child conditions assessed at the first home visit might have changed at the point of exit. Note that these additional questions were not asked of children who dropped out of the study (because the interviewers would not have known that the previous visit was going to be the final exit visit).

All questionnaires were administered using the Computer-Assisted Personal Interviewing software CSPro (Version 7.1.3), and OTP and Ration cards were scanned and data entered digitally using the SurveyCTO software. Questionnaires were translated into Hausa and administered to all respondents in Hausa.

Sample size (intended)

The sample size determination was based on the relapse rate as the principal outcome variable.

The minimum sample size required for this study was calculated to be 500 OTP-cured and 500 community control children across the 25 facilities. This sample size would allow us to detect a 4% point difference in SAM incidence between both cohorts of children with 95% confidence. Calculation parameters were chosen conservatively and assumed an incidence of SAM among community controls of 1%, a total number of 25 clusters (health facilities), a coefficient of variation of cluster sizes of 0.9, and an intra-cluster correlation of 0.02. Sample size calculations were implemented using the *clustersampsi* tool in Stata.

With an anticipated loss to follow-up of 20%, the study therefore aimed to recruit 600 children per cohort.

Note that children were not sampled. Rather:

1. All OTP-cured children and their caregivers were approached to be included in the study as they were being discharged from the CMAM programme (during our recruitment period) and assessed for eligibility until a maximum number of participants was recruited.

2. Community control children were purposively recruited into the study within the communities of each OTP-cured child using a snowball approach.

Sample size (actual)

- Out of a total of 645 OTP-cured children that were recruited into the study, 553 were found at the first home visit and deemed eligible, 83 were not found and 9 were later discovered to not be eligible.

- Out of a total of 543 community children that were recruited into the study (met eligibility criteria and were matched to OTP-cured children) at the first home visit, 17 were discovered later on to not have been eligible (and therefore were dropped).

  ➔ Therefore, the sample size of the study (and consequently the number of observations of the published data) consists of 553 OTP-cured children and 526 community control children.

In terms of the outcomes of children at the end of the study:

- Of the 553 OTP-cured children, 378 did not experience SAM during the study period and therefore lasted until the 12th home visit; 134 experienced SAM during one of the visits (at which point their inclusion in further follow-up visits ended); 32 dropped out

of the study at some point (3 withdrew their consent later, and 29 were no longer traceable); and 9 died during the study period.

- Of the 526 community control children, 488 did not experience SAM during the study period and therefore lasted until the 12th home visit; 3 experienced SAM during one of the visits (at which point their inclusion in further follow-up visits ended); 30 dropped out of the study at some point (6 withdrew their consent later, and 24 were no longer traceable); and 5 died during the study period.

Figure 2: Overview of study sample – OTP-cured children

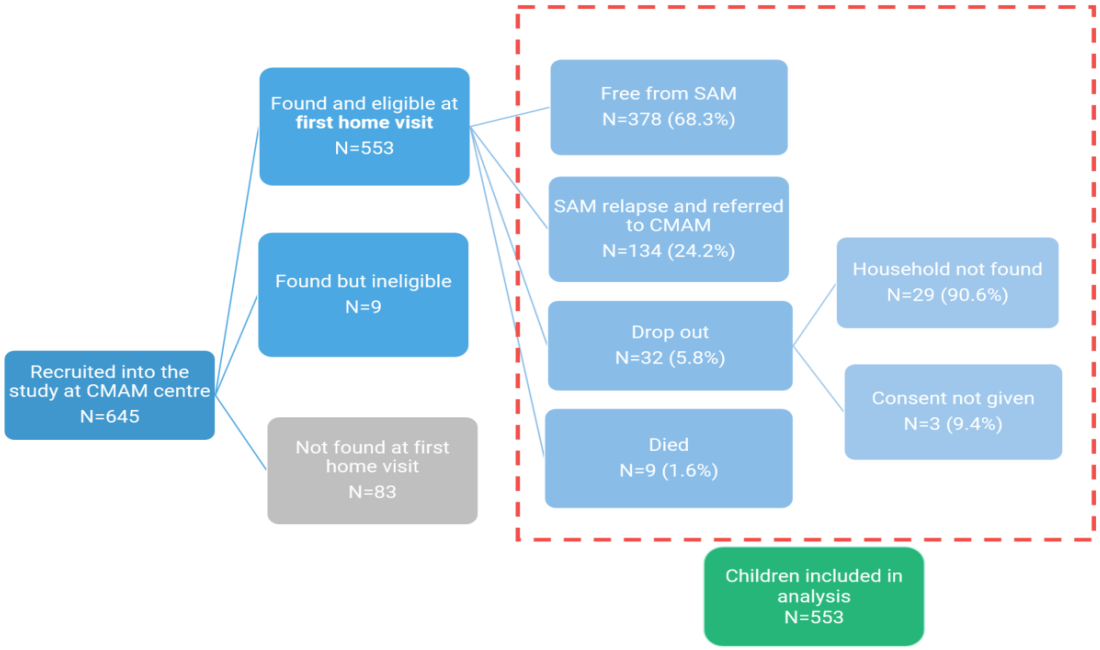
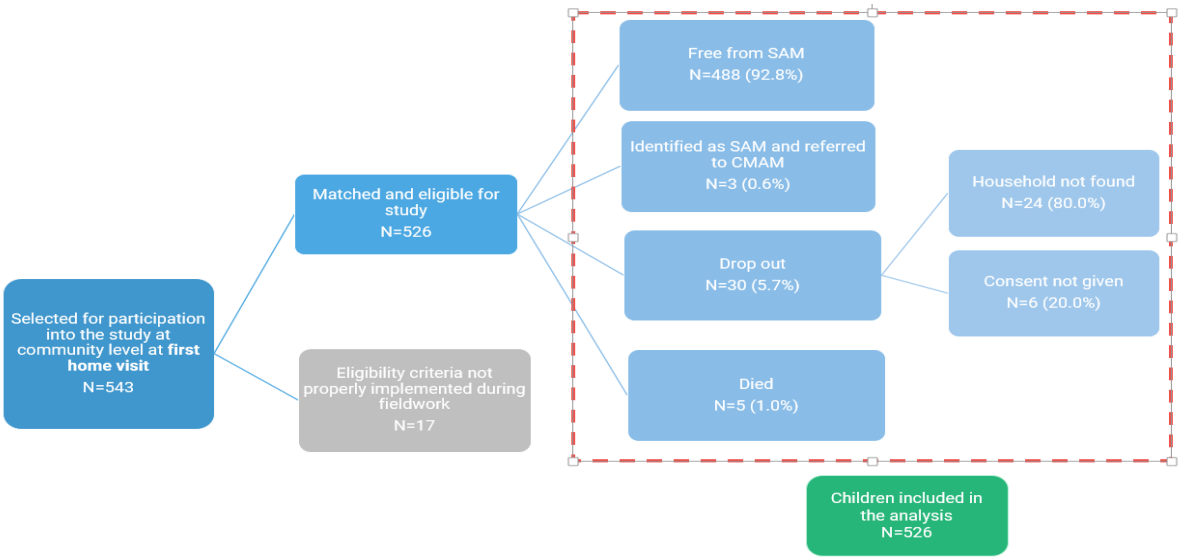


Figure 3: Overview of study sample – community controls



Ethical considerations

This study met the ethics criteria of the Sokoto State Health Research Ethical Committee and approval was received on 12 March 2018.

Verbal informed consent was sought and recorded in the questionnaire from mothers during recruitment of children into the study and from household heads at the first home visit. Consent was sought again from households at each follow-up visit. All children who experienced SAM during the study were referred to the nearest OTP services.

Training and data quality assurance

Before data collection was conducted, training was organised for the interviewing teams and included a mix of in-class training and piloting. Two trainings were organised: one for the recruitment phase and one for the home visits phase. The main objective of the training was to ensure that the data collection teams mastered the questionnaires, could measure MUAC accurately and implement the survey protocol, and were comfortable using CAPI.

Classroom training for the recruitment phase was structured following the recruitment and health facility questionnaire: for each module a brief introduction was delivered, then each module question explained, and finally a mock interview between trainees took place. The training ended with a two-day piloting exercise to practice using the instruments and protocols.

Similarly, classroom training for the home visits phase was structured to follow the long household questionnaire: for each module a brief introduction was delivered, then each module question explained, and finally a mock interview between trainees took place. A full day of training was dedicated to MUAC measurement and young children from Sokoto were invited for in-class practice. The training ended with a two-day piloting exercise to further familiarise trainees with overall survey instruments and protocols.

A central component of quality assurance was the supervision and feedback that each enumerator received during the training, piloting, and roll-out of the study. At the beginning of every training day, the trainees had to complete a test on the modules covered the previous day and individual feedback was provided daily to identify and resolve any challenges faced by the interviewing team.

Team supervisors were selected from among the most experienced and best-performing participants, and these individuals completed an additional training module for the extra tasks of coordination and quality assurance.

Several data quality assurance mechanisms were implemented to ensure data quality throughout the survey:

1. Data was collected using CAPI, which enabled automated live data checks during implementation of interviews. Extensive validations and cross-checks were programmed into the CAPI software and pre-tested to reduce errors and inaccuracies during the interviews.

2. Data were uploaded every day, which enabled the survey management team to conduct a range of consistency checks on a daily basis. Any issues identified at this stage were immediately communicated to the relevant team supervisors for action.

3. A data collection monitoring dashboard on PowerBi was used to daily monitor the progress of data collection as well as the performance of data collection teams and individual enumerators, thus allowing the field management team and team supervisors to give feedback to teams on a regular basis and continuously improve the quality of data collection.

4. Two measures of MUAC were taken at each visit to minimise the risk of measurement error. A third measure was triggered whenever the difference between the two first measures was more than 5mm.

5. Team supervisors were trained in quality assurance to control the quality of data collection in their teams and give live feedback to their team members.

6. The survey management team visited data collection teams at random throughout the implementation of the study to observe interviews and provide additional feedback to team supervisors and teams in general.

Definition of variables

SAM was determined using the WHO and national MUAC criteria of MUAC <115 mm. Given that this study's objective was to identify definite relapses and cases of SAM that would require treatment, **we classify a child as having SAM if his/her MUAC ≤112 mm at any home visit or if his/her MUAC is between 112 and 115mm for two consecutive visits.** The reason we do this is to account for the possibility of measurement error, i.e. it is difficult to identify whether children around the 115mm MUAC cut-off temporarily dip into SAM or whether they are a certain SAM case that requires treatment.

Stunting was defined using the WHO methodology and reference tables. A child was classified as stunted if his/her LAZ/HAZ was <-2 standard deviations (SD) away from the WHO reference median. A child is considered as severely stunted if his/her LAZ/HAZ was <-3 SD from the WHO reference median.

Notes on the data collection that might be relevant for data analysis

Survival analysis techniques can be used to analyse the data.

An important point to emphasise is that for OTP-cured children there was a lag of up to three weeks between their recruitment at the health facility and the first visit at home (where we collected baseline data). Some children had already relapsed into SAM at the first home visit before additional data on these children could be collected. For those children who relapsed between recruitment and the first home visit, it would therefore not be possible to assess whether certain time-varying characteristics collected at the first home visit – e.g. child-level health indicators – materialised as a consequence of relapse or prior to relapse. This is a limitation and could present implications for data analysis, depending on the type of analysis the users of the data wish to conduct. Specifically, it is important for the analysis not to suffer from endogeneity if for instance users are interested in assessing the effect of certain factors on relapse rates. There are options to

deal with this limitation, for instance, i) limiting the analysis to the factors/covariates that could reasonably be assumed to be time-invariant between recruitment and the first home visit, or ii) defining the time origin for OTP-cured children as the first home visit (as opposed to their recruitment from the health facility) and restricting the analysis to the subsample of children that had not relapsed into SAM at the first home visit (though this option would entail a significant reduction in sample size).

It is also important to emphasise that the two cohorts of children included in this study are not necessarily representative of the overall population of children in Northern Nigeria or even Sokoto State. On the one hand, children from the OTP-cured cohort were recruited from health facilities in a purposefully selected set of LGAs within that state. Given the way that they were recruited, they do represent a census of OTP-cured children from those health facilities that were discharged as cured during the recruitment phase of this study and that were eligible for the study. However, the level of representativeness beyond that group is unclear. Community control children, on the other hand, were selected using snowball sampling, which essentially implies purposeful sampling within visited communities. Hence, generalising findings beyond the two groups covered in the study should only be done with care.

Finally, we note for the user that the **collected data does not have any sampling weights**. This is due to the nature of child selection, whereby all children that were discharged as cured during the study's recruitment period and were eligible were invited to participate in the study.

Description of datasets

The uploaded data includes five datasets (in Stata format) that are described below. In addition to the datasets, we have also uploaded the survey questionnaires and we advise users to refer to the questionnaires when working with the data. All variables in the datasets can be easily linked to the corresponding questions in the questionnaire as the variable names match the question number in the questionnaire.

Note that the datasets and questionnaires for the recruitment of the OTP-cured children and community control children have not been included (these questionnaires solely assessed the eligibility of the children to be recruited into the study). However, relevant indicators from these questionnaires have been included in the ***'baseline_main'*** dataset. These are the sex of the child (*ch_gender*), the age of the child in months (*ch_age*), the level of education of the mother (*ma_education*), and the date of the recruitment of the OTP-cured child which is their date of discharge from the health facility (*n_date_recruit*).

Similarly, note that the data from the scanned OTP and Ration cards for the OTP-cured children has not been uploaded for public use due to data quality concerns. The data suffers from many missing observations, given that this data was not directly collected by the enumerators but relied on health facility staff filling in the OTP and Ration cards for the treated children. Only one variable from these cards is included in the uploaded data and this is the MUAC of the child at the time of their admission into OTP. This is represented by variable '*muac_entry*' which is included in the ***'baseline_main'*** dataset.

'baseline_main' dataset

This dataset corresponds to the **baseline questionnaire**, which was the long questionnaire administered at the first home visit for both OTP-cured and community control children. The

dataset contains data at the level of the household, as well as data at the level of the recruited child and their mother (which is also at the level of the household given that only one child per household was selected). The dataset includes the data from all modules of the baseline questionnaire except the household roster which is Section B of the household questionnaire.

This dataset has been anonymised for confidentiality purposes and therefore all identifying information have been removed from the data.

The unique identifier of this dataset is child_id. This dataset can be linked to all other datasets using the unique child ID (child_id). It can also be linked to the health facility dataset using the facility ID (hf_id).

There are some non-response observations in the data. A response of 'don't know' is coded as '*999*' in the dataset; a missing response (i.e. a response should have been provided but was not) is denoted as *'.a'* in the dataset; and a skipped response (i.e. a valid non-response due to a skip in the questionnaire) is denoted as *'.'* in the dataset.

Note that in addition to the questions in the baseline questionnaire, the *baseline_main* dataset also includes selected constructed indicators prefixed by n_ (these were constructed by data analysts using the survey data). These constructed indicators can be used to set up a survival analysis model and are included in the dataset to save data users time as they might require complex reshaping of data (but they could be generated by data users if preferred). The constructed variables include the following:

- The date of the OTP-cured child's recruitment into the study (on the day of their discharge from the health facility): *n_date_recruit*
- The date of the baseline questionnaire for the OTP-cured child and the community control child (for the community control child this would represent their date of recruitment into the study): *n_date_baseline*
- A dummy variable to indicate if the child developed SAM at any point during the 6 months follow-up, or if the child did not develop SAM (either lasted until the end of the study SAM-free, or dropped out of the study earlier): *n_event_outcome*
- A dummy variable to indicate if the child died during the 6 months follow-up: *n_died*
- A dummy variable to indicate if the child dropped out during the 6 months follow-up (household no longer consented to continue the study, or household moved away, etc.): *n_dropout*
- A dummy variable to indicate if the child lasted until the end of the study period SAM-free: *n_end_study*
- The number of the round (or visit) when the child exited the study (this would be the visit when the child developed SAM, or the final 12th visit if the child remained SAM-free, or the round before the child dropped out of the study and could no longer be traced): *n_round_exit*
- The date of the interview at the round (or visit) when the child exited the study: *n_date_followupexit*

- A variable indicating the number of days that elapsed from the date of joining the study until the outcome event (i.e. until either developing SAM, or dropping out, or lasting until the end of the study SAM-free): *n_days*

'baseline_household_roster' dataset

This dataset corresponds to the **baseline questionnaire**, which was the long questionnaire administered at the first home visit for both OTP-cured and community control children. The dataset contains data at the level of the household member and includes the data from the household roster which is Section B of the household questionnaire.

This dataset has been anonymised for confidentiality purposes and therefore all identifying information have been removed from the data.

The unique identifier of this dataset is child_id + serial_number. This dataset can be linked to all other datasets using the unique child ID (child_id).

There are some non-response observations in the data. A response of 'don't know' is coded as '*999*' in the dataset; a missing response (i.e. a response should have been provided but was not) is denoted as *'.a'* in the dataset; and a skipped response (i.e. a valid non-response due to a skip in the questionnaire) is denoted as *'.'* in the dataset.

'baseline_reproductive_history' dataset

This dataset corresponds to the **baseline questionnaire**, which was the long questionnaire administered at the first home visit for both OTP-cured and community control children. The dataset contains data at the level of the mother's births and includes the data from the reproductive history roster which is Section B of the mother questionnaire.

This dataset has been anonymised for confidentiality purposes and therefore all identifying information have been removed from the data.

The unique identifier of this dataset is child_id + mb_10. This dataset can be linked to all other datasets using the unique child ID (child_id).

There are some non-response observations in the data. A response of 'don't know' is coded as '*999*' in the dataset; a missing response (i.e. a response should have been provided but was not) is denoted as *'.a'* in the dataset; and a skipped response (i.e. a valid non-response due to a skip in the questionnaire) is denoted as *'.'* in the dataset.

'followups' dataset

This dataset corresponds to the **follow-up questionnaire**, which was administered at each follow-up home visit after the first (baseline) home visit for both OTP-cured and community control children. All follow-up visits for each recruited child are included in this dataset. The dataset contains data at the level of the household, as well as data at the level of the recruited child and their mother (which is also at the level of the household given that only child per household was selected). The dataset includes the data from all modules of the follow-up questionnaire.

This dataset has been anonymised for confidentiality purposes and therefore all identifying information have been removed from the data.

The unique identifier of this dataset is child_id + round_fu. This dataset can be linked to all other datasets using the unique child ID (child_id). It can also be linked to the health facility dataset using the facility ID (hf_id).

There are some non-response observations in the data. A response of 'don't know' is coded as '*999*' in the dataset; a missing response (i.e. a response should have been provided but was not) is denoted as *'.a'* in the dataset; and a skipped response (i.e. a valid non-response due to a skip in the questionnaire) is denoted as *'.'* in the dataset.

Note that not all modules from this questionnaire were administered at every follow-up. As mentioned in the protocol above, at every follow-up the child's MUAC was measured and data was collected on illnesses they may have experienced in the preceding two weeks and their breastfeeding status. However, at the exit round/visit (which is either the visit when the child has developed SAM or the final 12th visit if the child had remained SAM-free), additional modules were administered to the mother of the child to collect data on mother's employment, deaths in the household, household food security, mother's knowledge on feeding, mother's networks, mother's reproductive history, and child and household dietary diversity. The follow-up questionnaire indicates which modules and questions were asked at every follow-up and which were only asked at the exit round. Note that for the children that dropped out of the study either because household withdrew consent for participation or because household moved away, the exit round modules were not administered (that is because the enumerator did not know that the last round they visited the household was going to be the final one).

'health_facility' dataset

This dataset corresponds to the **health facility questionnaire**, which was administered once in each health facility on the first day the interview team visited the health facility to recruit OTP-cured children. The dataset contains data at the level of the health facility and includes data from all modules of the health facility questionnaire.

This dataset has been anonymised for confidentiality purposes and therefore all identifying information have been removed from the data.

The unique identifier of this dataset is hf_id. This dataset can be linked to all other datasets using the health facility ID (hf_id).

There are some non-response observations in the data. A response of 'don't know' is coded as '*999*' in the dataset; a missing response (i.e. a response should have been provided but was not) is denoted as *'.a'* in the dataset; and a skipped response (i.e. a valid non-response due to a skip in the questionnaire) is denoted as *'.'* in the dataset.