Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Version 07.0 (28.05.2015), includes 8 appendixes

Prepared by:

Lotta Alho, MSc, study statistician

Yin Bun Cheung, PhD, responsible study statistician

Jan Peerson, MSc, study statistician

Approved by:

Name, degree	**Position**	**Signature**
Kenneth Maleta, MBBS, PhD	Principal investigator	
Per Ashorn, MD, PhD	Co-Principal investigator	
Ulla Ashorn, PhD	Responsible investigator, social studies	
Kathryn Dewey, PhD	Principal investigator for the iLiNS-Project	
Steve Vosti, PhD	Responsible investigator, economic studies	

Contact details: Ms. Lotta Alho, University of Tampere School of Medicine, Finland. Email: Lotta.Alho@uta.fi

Table of Contents

1 Version history

Version number	Version date	Prepared by	Description of the completed editions
01.0	17.03.2013	Lotta Alho Yin Bun Cheung Jan Peerson	Original document (includes appendix 01)
0.20	07.05.2013	Alho, Cheung, Peerson Beth Prado, Chiza Kumwenda	Added Appendix 02: Developmental outcomes at age 18 months (prepared by Beth Prado) and Appendix 03: Breast milk intakes (prepared by Chiza Kumwenda)
03.0	24.05.2013	Alho, Cheung, Peerson Anna Pulakka	Added Appendix 04: The effect of LNS on physical activity (prepared by Anna Pulakka)
04.0	26.07.2013	Alho, Cheung, Peerson Mary Arimond	Added Appendix 05: Effect on infant and young child feeding practices (prepared by Mary Arimond)
05.0	25.04.2014	Alho, Cheung, Peerson Aleks Schaefer	Main growth SAP: Edited the way SAE's are claculated and presented. Removed adjustment for time in the study and presenting total number of cases per group. Added Appendix 6: Hypothetical Willingness-to-Pay for LNS and Likuni Phala at Baseline (Prepared by Aleks Schaefer)
06.0	28.04.2014	Alho, Cheung, Peerson Mary Arimond	Revised Appendix 05 into version 2.0: Effect on infant and young child feeding practices (prepared by Mary Arimond)
07.0	28.05.2014	Alho, Cheung, Peerson Ulla Ashorn	Added Appendix 7: Long-term acceptability of LNS (lipid-based nutrient supplements) for infants in Malawi (Prepared by Ulla Ashorn) Added Appendix 8: The effect of Lipid-based

		Jaden Bendabenda	Nutrient Supplement (LNS) provision on child morbidity (Prepared by Jaden Bendabenda)

2 Introduction

Poor growth and severe childhood stunting are very common in rural Malawi and elsewhere in Sub-Sahara Africa, with known negative consequences for child development and long-term individual and household welfare. To date, few interventions have proven successful in promoting linear growth in early childhood. Preliminary results from Malawi and Ghana suggest that a 6-12 month-long daily complementary feeding of infants with 20-50 g of an energy-dense and highly micronutrient fortified Lipid-based Nutrient Supplement (LNS) may markedly reduce the incidence of severe stunting before the age of 18 months.

The iLiNS-DOSE trial was designed to study the impact of a 12-month LNS provision to infants and young children on their growth, nutritional status and a number of other health outcomes. For this purpose, a total of 1932 infants were enrolled in rural Malawi, randomized to receive for 12 months either no supplement or one of five alternative LNS-preparations (six study groups in total), and intensely monitored for 12-months for a large number of outcomes. Key details of the trial have been recorded at the clinical trial registry at the National Institutes of Health (USA) (www.clinicaltrials.gov), under the registration number NCT00945698. A full trial protocol is available from the contact person to this document.

This document (called "the statistical analysis plan" or SAP) describes the study group's plan for data analysis, management, and storage. The SAP is designed to be evolving over time. Version 1.0 documents the details of the hypothesis testing and other analyses on primary and selected secondary outcomes. Subsequent versions of the SAP will give further details on the analyses and hypothesis testing of additional secondary outcome variables and exploratory analyses from the data.

3 Study objectives

The main aim of the trial is to assess the safety and impact of LNS supplementation on linear growth of infants and young children and to identify individual, household, and village-level characteristics that would modify the effects of LNS on child growth. A secondary aim is to similarly study the impact of LNS on various other (secondary) outcomes in the same target group. Finally, the trial aims to provide descriptive information on issues that might be necessary to facilitate future demand creation for LNS interventions.

The above aims have been broken down into the following first five objectives that were predefined in the trial protocol. The safety aim was not explicitly stated among the predefined objectives in the trial protocol, but was listed under the primary outcomes for analysis.

1. To determine the lowest daily dose of LNS that will promote linear growth among 6-18 months old infants in a rural Malawian community with poor food security. *This covers also the objective to study if ANY dose of LNS promotes linear growth among the target group.*

2. To assess if LNS made without milk has a linear growth-promoting effect comparable to that of milk-containing LNS (in the above described environment, with 20 and 40 g/day ration sizes).

3. To determine the impact of five different LNS supplementation schemes on child dietary intake, morbidity, appetite, micronutrient status, immune function, and neuro-behavioral development.

4. To collect information to facilitate future demand creation for LNS interventions, such as end-user knowledge, attitudes and practices related to LNS and other feeding and parental care-giving practices.

5. To assess the extent to which household food insecurity and other individual, household, and village-level characteristics modify the effects of LNS on child outcomes.

6. To determine if any of the LNS supplements lead to increased risk of serious adverse events.

4 General approach to data analysis

There will be three categories of data analysis.

1. For the main aim (safety and linear growth outcomes), the analyses will be driven by predefined primary study hypotheses (see chapter 4 below). Conclusions on this part of the study will be based on formal hypothesis testing.

2. For the secondary aim (other outcomes), the analyses will be driven by similar hypotheses to those used for linear growth. These hypotheses have not been predefined in the trial protocol and hence they do not appear in version 1.0 of this SAP. They will, however, be defined as appendixes in subsequent versions of the SAP. For each hypothesis-driven analysis, the SAP will be updated prior to starting the analysis.

3. In addition to the hypothesis-driven questions, there will be a large number of exploratory analyses. In the absence of predefined study hypotheses, these analyses will be considered hypothesis-generating, rather than confirmatory.

5 Hypotheses to be tested

As indicated above, version 1.0 of the SAP describes predefined hypotheses only for the linear growth outcomes, i.e. specific objectives 1 and 2. Further hypotheses will be formulated and documented in subsequent SAP versions before the respective analyses are started.

Objective 1 / hypothesis 1: The change in mean length-for-age Z-score (LAZ) of infants provided with 10, 20, or 40 g/day of LNS from 6 to 18 months of age will be greater than that of infants who receive no dietary intervention at the same age.

Objective 2 / hypothesis 2: The change in mean LAZ of infants receiving 20 or 40 g/day of LNS without milk from 6 to 18 months of age will not be lower than that of infants receiving a comparable intervention with milk-containing LNS.

6 Data cleaning and procedures on breaking the intervention code

The study group will adopt the following procedures for data cleaning and breaking the intervention code

1. In the first phase, a number of investigators will do preliminary cleaning of the data required to the main analyses (safety and linear growth outcomes). At this point, all investigators are totally blinded to the intervention each participant has been receiving.

2. A study statistician (L.A) makes a preliminary database that contains semi-clean data required for the main analyses. The database and summary statistics for each variable are distributed to the principal investigators, the members of the board governing trial implementation and the principal biostatistician for the trial. Once these individuals agree that the data are sufficiently comprehensive and clean, the study statisticians (L.A, J.P, and Y.B.C) are provided with the database and a code that can be used to group the participants who received the same intervention together – i.e. that gives group codes 1, 2, 3, 4, 5, and 6 without indicating the actual intervention each group number relates to.

3. The study statisticians review the data and complete preliminary analyses for group comparisons (without knowing the actual interventions). Based on these analyses, the study statisticians make suggestions for the amendment of the SAP (e.g. on the treatment of missing values). The investigators listed under 2) above then agree on a revised version of the SAP, after which the intervention code is broken and the main analyses are completed.

4. For secondary outcomes, the analyses will be mostly completed by investigators who are not study statisticians. For each of these analyses, data cleaning will be completed as above. Once the analyst has completed the first round of data cleaning without any knowledge about the group information, s/he will request scrambled group information from the

statisticians. This information will again group the participants who received the same intervention together without indicating the actual intervention each group number relates to. For each analyst, the study statisticians provide a new / different set of scrambled group codes – so that two analysts cannot combine their results during the analysis.

4. Before the intervention code is fully broken, mistakes found in the data can be corrected in the database, as long as there is an audit trail that indicates the date of correction, the old and new value, justification for the correction and the identity of the person authorizing the change (this is not necessary for the correction of entry errors). After the code is broken, the data on main outcomes will be "frozen" and data can no longer be corrected in the database. Instead, all corrections (also entry errors) will be reviewed and need to be approved by the responsible investigator and documented before programmed into cumulative syntax-files (do-files, one for each data collection form) that will contain the same information as the audit trail described above. These do-files need to be run to clean the data before any subsequent analyses.

5. Data cleaning for other data not used for the main analyses will continue even after breaking the intervention code. For each additional data collection form, the data will be similarly frozen by the time first real analyses will be completed from them (the time can vary form by form). Also for these forms, mistakes found before data freezing will be corrected straight into the database whereas those found after the data freezing will be corrected in separate data-cleaning do files. Both correction methods will contain the audit trail that can be used to track all completed changes.

6. Any investigator may raise a suspicion for a correctable mistake in the data. If such a suspicion arises, the investigator who has the responsibility over those particular data (each data collection form has a defined responsible investigator) should be informed and s/he should investigate if a correction is needed. If yes, the data managers in Finland and Malawi will be informed and the change will be made and documented either to the database (before data freezing, this will be done in Malawi), or to a correction do-file (after data freezing, this will be done in Finland).

7 Definition of the primary outcomes

Change in length-for-age Z-score (LAZ):

Length for age will be determined from age, sex, and length information at six months (visit 0) and 18 months (visit 52) of participant age, using the Stata macro developed by the WHO using the WHO 2006 multi-centre growth standard. The values will be rounded to two decimal points. The change (expressed with two decimals) will be calculated by subtracting the value at 6

months from the value at 18 months of age. *The data will be extracted from Form 00: Q2.5; Form 04: Q1.2, Q2.2*

8 Safety outcomes

Incidence of serious adverse events (SAE) during the observation period.

SAE will be defined as an event determined to be an SAE by the study physician. The SAEs will be categorized into five categories: Death, life threatening event, inpatient hospitalization, significant disability or other serious adverse event. *The data will be extracted from Form 29, Q3.2.*

9 Basis for the analysis: Intention to treat and per protocol

Primarily, the analysis will be based on the principle of modified intention-to-treat. The modification concerns six participants who were accidentally allocated to another group than actually randomized. For each participant, the randomization code was pre-packed and sealed in an individual envelope that was opened and used for group allocation at enrolment. For these 6 individuals, the randomizer made a recording error, i.e. s/he noted down in a data collection form an incorrect group code or wrote the code with unclear handwriting. The incorrect code was later transcribed into the computer software that was used to plan participant visits and allocate interventions. These six participants were told to belong to the erroneously recorded intervention group and they received that intervention throughout the trial – hence they will also be analyzed in that group (rather than the one written on the randomization slip).

All randomized participants are eligible to be included in the analyses, with the exception that subjects with missing data on an outcome variable will be excluded for the analysis of that outcome. For variables targeted to be measured every 6-months, the data are considered missing if the actual measurement date is over +/- 8 weeks from target.

Number of participants with non-missing values analyzed for each end point will be presented by treatment groups.

Besides the intention to treat, there will be two per protocol analyses, adjusting for the adherence to the dietary supplementation. In the first per protocol analysis, the statistical model will include a variable that indicates the number of intervention doses delivered to the participant during the time period for which the participant received the study supplements (the supplements were delivered every two weeks). In the second per protocol analysis, the statistical model will include a related variable that indicates the number of days when the participant's guardian indicated that the participant had eaten the study supplement (this information was collected on a daily basis).

10 Time points for the analyses

All the main analyses will primarily cover the period from enrolment (when the participants are 6 months old) to the end of the intervention (when the participants are 18 months old).

Secondarily, there will be a growth analysis after a 2-year post-intervention period, i.e. covering a period from enrolment to when the participants are 42 months old. SAE data will not be collected or analyzed after the participants have turned 18 months old.

11 Presentation of the study findings and hypothesis testing

11.1 Success of enrolment and follow-up

All registered participants and the success of their follow-up will be described in a flow chart (Figure 1). For additional information drop-out rate between groups will be tested with Fisher's exact test and baseline characteristics of drop-outs compared to those who completed the study will be tested with t-test or chi square. P-values for these tests will be shown in the text.

11.2 Baseline information

Participant characteristics at enrollment will be tabulated by treatment arms as indicated in tables 1a and 1b. Hypothesis testing will be performed for baseline information to give additional information but p-values will not be presented in Table 1 of the eventual manuscript. Methods used for hypothesis testing are indicated in Tables 1a and 1b.

11.3 Comparison of the change in length-for-age Z-score between intervention groups versus the control group

Box-whisker plots of the change in LAZ of all 6 groups will be shown side by side in Figure 2. The group means and standard deviations for the change in length (in cm) and LAZ (in z-score units) will be presented as indicated in Table 2. The table will also tabulate the difference in mean length and LAZ and their 95% confidence intervals between the control group and each of the intervention groups.

Mean change in LAZ in each of the intervention groups will be individually compared against that of the control group using a two-sample t-test. The Holm's method will be used to adjust for the multiple comparisons (Aickin & Gensler, 1996). The raw P-values are presented in tables but the text and conclusion will be based on the Holm-adjusted P-values.

The null-hypothesis of LNS having no impact on linear growth is rejected for each comparison that yields a Holm's adjusted p-value <0.05.

11.4 Comparison of milk-containing and milk-free LNS

This analysis will exclude participants who received no intervention or 10 g / day milk-containing LNS. An ordinary least-square regression of the following form will be estimated:

$$\hat{y}_i = b_0 + b_1 D_i + b_2 S_i + b_3 (D_i \times S_i)$$

where D= 0 or 1 for the 20g and 40g group, respectively, and S= 0 or 1 for the milk and milk-free LNS group, respectively. If the coefficient b3 for the interaction term does not reach a significance level of P<0.10, a simpler model without the interaction term will be estimated. 90% CI will be provided for each parameter. Ninety per cent instead of 95% CI is used because the non-inferiority consideration is one-sided and the use of 90% CI is the convention in this setting (Senn, 1997). Both models will be shown in table 3. Conclusion is to be based on the simpler model if the interaction term is not statistically significant (P>0.10), or vice versa.

Evaluation

The set non-inferiority margin for this analysis will be 0.15 Z-score units. If the lower bound of the 90% CI of the coefficient b2 is larger than -0.15, non-inferiority is confirmed. If the upper bound of the 90% CI is smaller than -0.15, inferiority is confirmed. If the upper and lower bounds of the 90% CI are larger and smaller, respectively, than -0.15, the finding is not conclusive.

If there is no significant interaction, the above evaluation only need be done once using the simple regression model without interaction. If there is significant interaction (P<0.10), the above evaluation will be done twice based on the model with interaction: Firstly on the 90% CI of the b2 coefficient concerning the non-inferiority of milk-free LNS in the 20g dose. Secondly on the 90% CI of the (b2+b3), obtained by the linearly combination command (Stata's **lincom**), concerning the non-inferiority of milk-free LNS in the 40g group.

11.5 Safety profile: Analysis of serious adverse events

The total number of serious adverse events (SAEs) will be presented by intervention group and SAE categorization, as indicated in Table 4. There will also be an analysis on the proportion of participants with at least one episode of SAE. For each of the intervention groups, the proportion and a relative risk (as compared to the control) will be presented as indicated in table 4. Fisher's exact test will be used to test the global null hypothesis of no differences between groups and the null hypothesis will be rejected if P<0.05. If the global null hypothesis is rejected, comparison between each intervention group and the control group will be conducted using log-binomial regression. Rate ratios (95% CI) for the incidence of any SAE will be calculated for each intervention group (as compared to the control group).

AE information was not collected after the participants had finished the intervention and completed the primary outcome study visit at 18 months of age.

Cumulative incidence curve for mortality will be presented graphically by intervention group as illustrated in Figure 3. The log-rank test will be used for testing global hypothesis and pair-wise comparisons. Hypothesis of equality of survivor functions is rejected if $P<0.05$. Hazard ratios and their confidence intervals will be estimated by the Cox regression model. The **sts** and **stcox** commands will be used.

12 General notes on statistical methods

12.1 Software

All analyses will be done in Stata version 12. The WHO 2006 multi-centre growth standard will be used for age-and-sex standardization of weight, length (height), weight-for-height, MUAC and head circumference.

12.2 Preparing anthropometric data for analysis

All the anthropometric measurements were completed in triplicate during each study visit. But for the analysis, the team will use the mean of the first two readings if they do not differ more than by a pre-specified tolerance limit. If they do, the third measurement will be compared with the first and second measurements and the pair of measurements that has smaller difference will be used to calculate the mean which will be used in analyses. If there are only one or two repeated measurements mean of those two will be used for the analyses.

The agreed tolerance limits between the first two measurements are:

1. length/height $\leq$ 0.5 cm
2. circumferences (head, MUAC) $\leq$ 0.5 cm
3. infant/child weight $\leq$ 0.1 kg
4. adult weight $\leq$ 0.1 kg
5. skinfold thickness $\leq$ 2.0 mm

The length, circumference and skinfold thickness measurements were recorded to last complete unit (mm). To account for the bias of always rounding the values a bit downwards, half a unit will be added to all length, circumference and skinfold thickness measurements prior to the analysis. This procedure is not done for weight measurements, since they were recorded accurately with electronic scales.

Missing growth values will be treated as missing, i.e. there will be no growth data imputation from the other data. Imputation is considered unreliable due to long time-interval between anthropometric measurements, at an age when growth faltering is common.

When measuring change in anthropometric values between two time intervals, the change will be treated as missing if there is a missing value in either of the respective time points.

12.3 Multiple comparisons

The study involves multiple objectives and therefore multiple sets of hypothesis. Statistically, the different sets of hypotheses are considered independent families of hypotheses. Statistical adjustment for multiple comparisons in one family of hypotheses does not need to consider the other families.

For efficacy analysis, each family consists of 5 hypotheses, each comparing an intervention group versus the control group. The Holm's adjustment method is used.

For safety analysis, it is preferable to err on the cautious side (Nauta, 2010). We began with testing the global null hypothesis of no difference between groups. If the global null hypothesis is rejected, raw P-values are used in the comparisons between intervention and control groups.

12.4 Confidence intervals

Regardless of results in hypothesis testing, the calculated ratios and differences in between-group comparisons will be complemented with confidence intervals (usually at 95% level but 90% for non-inferiority studies), for descriptive purpose. For the quantitative outcomes, confidence intervals will be based on t-test. For binary outcomes, the confidence intervals will be based on binomial distribution.

12.5 Interaction and effect modification

There will be two sets of test for interaction between the intervention group and selected other variables on their association with the primary outcome (change in length-for-age z-score). All tests will be done using the likelihood ratio test.

The first set of analyses will be hypothesis-driven and will include unambiguous predefined variables that could logically modify the effect of the nutritional intervention on infancy and early childhood growth. Variables included (as continuous variables where possible) in this analysis include:

1. The participant's baseline length-for-age (below / above median)
2. Maternal BMI at enrolment
3. Maternal height
4. Maternal age
5. Maternal parity
6. The participant's sex
7. Number of under-five year old children in the household
8. Cohabitation of the child's father with the rest of the family (yes / no)

9. Household food security

The second set of analyses will be exploratory in nature and will include variables that can be constructed in several ways or that cannot a priori be logically linked to an effect modification. Themes or variables included in this analysis include:

1. Maternal knowledge, attitudes, and practices around child nutrition
2. Household wealth

If a statistically significant interaction ($p<0.1$) is found, the outcome analysis will be completed as stratified by the respective predictor variable. Variables that show no interaction with the intervention group can be used as covariates in the main analysis.

12.6 Covariate adjustment

The final decision on the use of covariates in main analyses will be decided based on preliminary analyses on the final dataset that includes information on the clustering of participants in the same group but does not provide information on the actual intervention delivered to each group.

At this stage, four different regression models for change in LAZ will be constructed, first of these with no covariate adjustments and models 2-4 with various adjustments. If the point-estimates for the regression co-efficient for the intervention group differ in any of the models by more than 10% from that in the unadjusted analyses, the adjusted analysis will be primarily presented in the eventual publications. If all the differences are smaller than 10% (Maldonado & Greenland, 1993), the results in the tables will be shown without covariate adjustments, but model 4 results will be provided in the text as supplementary information.

The four models include:

1. No covariate adjustment
2. Adjustment for baseline LAZ-score
3. Adjustment for baseline WHZ-score
4. Adjustment for baseline LAZ-score and baseline WHZ-score and for any of the variables presented in tables 1a and 1b showing statistically significant association (at $p<0.1$ level) with the change in LAZ-score

13 Storage and release of data

The data meta-data will be stored in a tailor-made hierarchical database, consisting of a MS Access front-end and MySQL tables in the back-end. The database, associated metadata, and form –specific do-files that contain all cumulative data corrections for the respective data collection forms are stored at a computer server at the University of Tampere and daily copied to

a server at the Mangochi research site, University of Malawi. A study statistician (L.A) acts as the manager for these data.

When an investigator wishes to perform certain analyses, s/he will request the respective data from the above-indicated data manager. The data manager will export all the data from the respective data collection form into an excel or Stata file, run the cumulative data correction do-file and then provide the corrected data, together with the syntax for the correction do file (that documents all the completed data editions) to the person requesting the data.

The databases and the do-files will be named with systematic naming format and stored at the central server at the University of Tampere. For each article, the following files will be stored:

1. The database from which the analyses were performed
2. The data dictionary
3. The data correction do file(s)
4. The data analysis do file(s)
5. The actual scientific article

The data collection forms and respective user guides will be stored at the central study repository, in the computer server at the University of Tampere

In the longer run, there is a plan to place the data publicly available in the internet.

13.1 Data and output handling

To ensure reproducibility and to keep an audit trail, all data management, analysis and outputting procedures will be kept as Stata **do** files. All transformation, categorisation, or creation of variables as well as keeping or dropping of subjects in specific analyses will be written in the do files. The **do** files are to be executed in order to obtain these new data features temporarily, as opposed to saving these new features into permanent data files. It is envisaged that a large number of commands are required, and they may need to be partitioned in more than one **do** file. Numeric values will be used to indicate the correct sequence for running these files and version number of the do file is indicated at the file name, e.g. iLiNS-DOSE data cleaning01, form 18, v01.0, 2012-12-27.do should be executed before iLiNS-DOSE data analysis02, form 18, v01.0, 2012-12-27. If data from more than one form are used the form number is not indicated in the do-file name but forms are listed in the comments section in the beginning of the do-file. Variables on data version and version date are included in the data file and people using the data are asked not to share the files with other approved data users. All approved users obtain the data from the data manager so that the latest version is distributed. Outputs will be saved as log files.

A master **do** file, for example, may include, but is not limited to, the following commands to execute all the data modification, analyses and outputting procedures in one go:

**** Example of a master do file

```
**** DOSE main paper, master do file

clear

version 11.2

set more off

set mem 50m

cd c:\dose\mainpaper

capture log close

log using mainpaper.log, text replace

do iLiNS-DOSE data cleaning01, form 18, v01.0, 2012-12-27.do

do iLiNS-DOSE data analysis02, form 18, v01.0, 2012-12-27.do

do iLiNS-DOSE data analysis03, form 18, v01.0, 2012-12-27.do

log close
```

14 Procedures and history on modifications to the analysis plan

All new versions of and additions to the statistical plan will be approved by a team of core investigators, consisting of the senior researchers who oversee the trial implementation (iLiNS-Malawi Board of Directors) and the study statisticians. Each version will be identified with a new version number and a date of approval and named with standardized file-name format (iLiNS-DOSE analysis plan, version 00.3, 2012-12-27.docx).

In the file name, the first two digits before the decimal indicate an approved change to the SAP (ie version 01.0 denotes the first approved version, 03.0 the third approved version etc). The last digit after the decimal indicates a yet unapproved revision number for a document under editions (eg. 02.1 points to a document that is based on the second approved version, but has undergone one round of yet unapproved editions to it).

The table below lists the editions made to the different approved versions of the SAP:

Version number	**Date of approval**	**Editions**
01.0	xx.xx.2013	Original document

15 References

Aickin M, Gensler H. Adjusting for multiple testing when reporting research Results: The Bonferroni vs Holm Methods. *American Journal of Public Health* 1996; 86(5): 726-728.

Maldonado G, Greenland S. Simulation Study of Confounder-Selection Strategies. *American Journal of Epidemiology* 1993; 138(11): 923-936.

Nauta J. *Statistics in Clinical Vaccine Trials*. Heidelberg. Springer. 2010.

Senn S. *Statistical Issues in Drug Development*. Chichester, GBR. John Wiley & Sons. 1997.

16 Appendixes

Statistical Analysis Plan, Appendix 1: Secondary growth outcomes

Statistical Analysis Plan, Appendix 2: Developmental outcomes at age 18 months (added on 07.05.2013).

Statistical Analysis Plan, Appendix 3: Breast milk intakes (added on 07.05.2013).

Statistical Analysis Plan, Appendix 4: The effect of LNS on physical activity (added on 24.05.2013).

Statistical Analysis Plan, Appendix 5: Effect on infant and young child feeding practices (added on 26.07.2013, revised into version 02 on 28.04.2014).

Statistical Analysis Plan, Appendix 6: Hypothetical Willingness-to-Pay for LNS and Likuni Phala at Baseline (added on 25.04.2014).

Statistical Analysis Plan, Appendix 7: Long-term acceptability of LNS (lipid-based nutrient supplements) for infants in Malawi (added on 28.05.2014)

Statistical Analysis Plan, Appendix 8: The effect of Lipid-based Nutrient Supplement (LNS) provision on child morbidity (added on 28.05.2014)

17 Legends to the figures

Figure 1. Participant flow in CONSORT recommended format (Lancet 2001: 357: 1193)

Figure 2. Box-Whisker plots of change in LAZ by groups

Figure 3. Cumulative incidence curve for mortality, by intervention groups

18 Figures

Figure 1. Participant flow

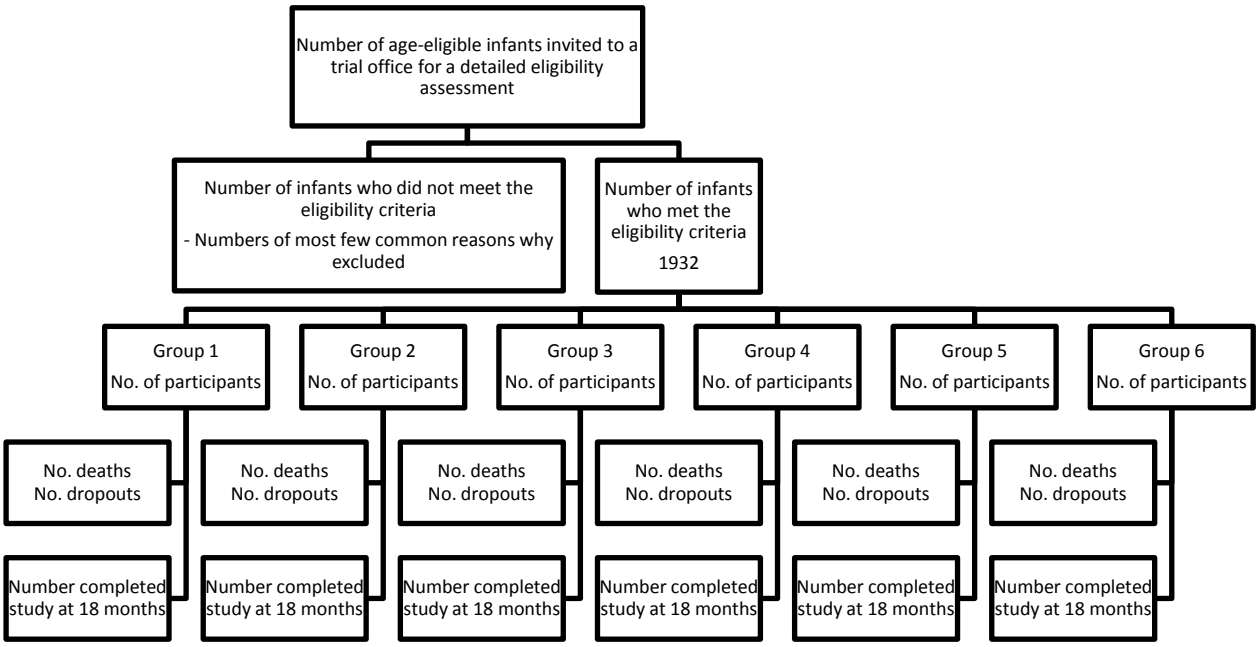


Figure 2. Box-Whisker plots of change in LAZ by groups

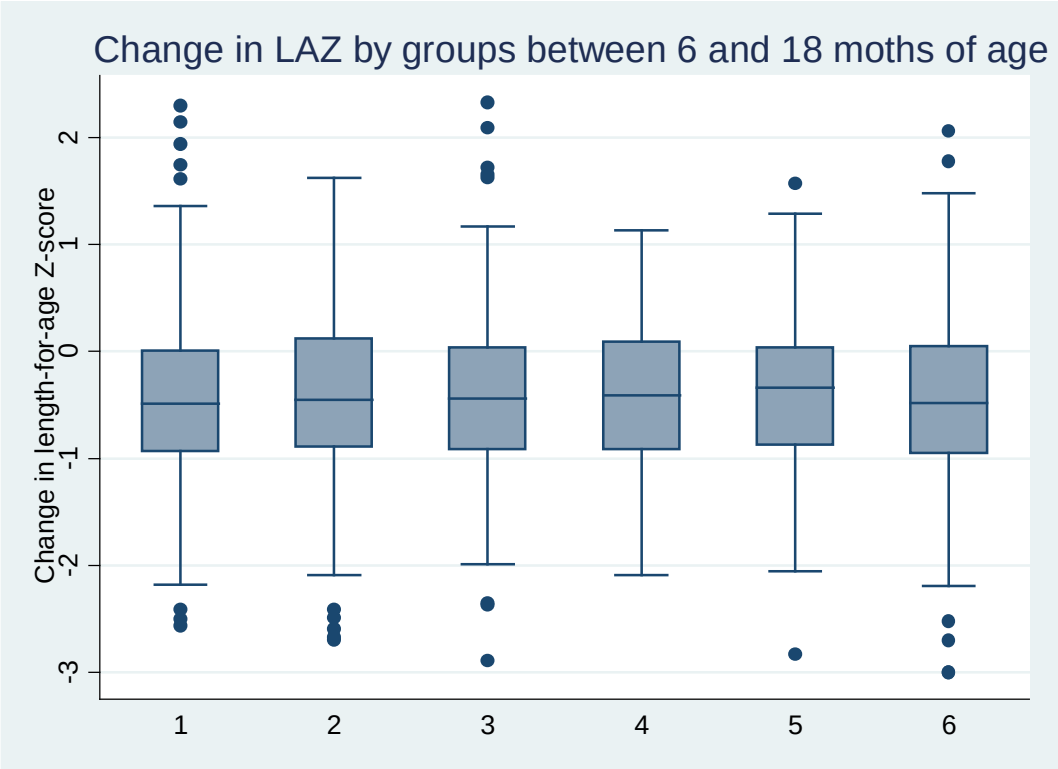


Figure 3. Cumulative incidence curve for mortality, by intervention groups

19 Tables

Table 1a Baseline characteristics of participants at enrolment

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**Test**
Number of participants	xxx	xxx	xxx	xxx	xxx	xxx	
Age, months (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Proportion of Males (percentage)	xxx / xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	Chi-squared
Weight, kg (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Length, cm (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Middle upper arm circumference, cm (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Head circumference, cm (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Weight-for-age z-score (mean, SD, N)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Length-for-age z-score (mean, SD, N)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Weight-for-length z-score (mean,	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA

SD, N)							
MUAC z-score (mean, SD, N)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Head circumference z-score (mean, SD, N)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Incidence of severe stunting at 6 months of age (percentage)	xxx / xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	Fisher's exact test
Incidence of moderate and severe stunting at 6 months of age (percentage)	xxx / xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	Fisher's exact test
Blood hemoglobin concentration g/dl (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Proportion of participants with anaemia (percentage)	xxx / xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	Chi-squared
ZPP concentration µg/g (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
ZPP concentration >45 (percentage)	xxx / xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	Chi-squared

Table 1b Other baseline characteristics to be assessed

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**Test**
Maternal Age, years (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Maternal height, cm (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Maternal Weight ,kg (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Maternal BMI (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Maternal MUAC, cm (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Maternal education, years (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Paternal age, years (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Paternal education, years (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Total number of persons in the household (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	Kruskal-Wallis test
Total number of children below 5 years of age in the household (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	Kruskal-Wallis test
Proportion with malaria parasiteamia at enrolment, number (percentage)	xxx / xxx (xx%)	xxx / xxx (xx%)	xxx / xxx (xx%)	xxx / xxx (xx%)	xxx / xxx (xx%)	xxx / xxx (xx%)	Chi-squared

Table 2 Anthropometric outcome changes during and up to 12 months of intervention

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**
Mean (SD) change in length-for-age z-score mean, Z-score units	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference (95% CI) in means between the indicated intervention and the control group		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Mean (SD) change in length, cm	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference (95% CI) in means between the indicated intervention and the control group		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)

Table 3 Regression results of the effects by the milk-containing and milk-free LNS

	Model: $\hat{y}_i = b_0 + b_1D_i + b_2S_i$			**Model:** $\hat{y}_i = b_0 + b_1D_i + b_2S_i + b_3(D_i \times S_i)$		
Variable	**Coefficient**	**90 % CI**	**P-value**	**Coefficient**	**90 % CI**	**P-value**
Constant (b_0)	xx.xx	xx.xx to xx.xx	x.xxx	xx.xx	xx.xx to xx.xx	x.xxx
Difference in mean change in LAZ in 40g dose LNS compared to 20g dose LNS (b_1)	xx.xx	xx.xx to xx.xx	x.xxx	xx.xx	xx.xx to xx.xx	x.xxx
Difference in mean change in LAZ in milk-free LNS compared to milk-containing LNS in 20 g group (b_2)	xx.xx	xx.xx to xx.xx	x.xxx	xx.xx	xx.xx to xx.xx	x.xxx
Difference in mean change in LAZ in milk-free LNS compared to milk-containing LNS in 40 g group ($b_2 + b_3$)				xx.xx	xx.xx to xx.xx	x.xxx

Table 4 Incidence of SAEs by study group

	Intervention group						
Intervention group	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**P-value**
Number of participants	xxx	xxx	xxx	xxx	xxx	xxx	
Number of participants with at least one episode of SAE (% with SAE)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx.xxx
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	
Risk ratio (95% CI) between the indicated intervention and the control group	1.00 (ref)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	
Number of participants with fatal event (% with SAE)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx.xxx
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	
Risk ratio (95% CI) between the indicated intervention and the control group	1.00 (ref)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	
Number of participants with life threatening event (% with SAE)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx.xxx
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	

Risk ratio (95% CI) between the indicated intervention and the control group	1.00 (ref)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	
Number of participants with hospitalization (% with SAE)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx.xxx
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	
Risk ratio (95% CI) between the indicated intervention and the control group	1.00 (ref)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	
Number of participants with significant disability (% with SAE)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx.xxx
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	
Risk ratio (95% CI) between the indicated intervention and the control group	1.00 (ref)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	
Number of participants with other SAE (% with SAE)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx.xxx
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	
Risk ratio (95% CI) between the indicated intervention and the control group	1.00 (ref)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)	

Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Appendix 01: Appendix 01: Secondary growth outcomes (added on 13.03.2013)

Contents

1. Study objectives

The main aim of the trial was to assess the safety and impact of LNS supplementation on linear growth of infants and young children and to identify individual, household, and village-level characteristics that would modify the effects of LNS on child growth. A secondary aim is to similarly study the impact of LNS on various other (secondary) outcomes in the same target group.

The aim of these secondary analyses is to compare the following outcomes between infants in 6 different intervention groups: dietary supplementation with 10 g, 20 g or 40 g /day milk-containing LNS, 20 g or 40 g /day milk-free LNS or nothing (delayed intervention).

1. Change in weight-for-age (WAZ), weight-for-length (WHZ), mid-upper arm circumference (MUAC)-for-age and head circumference-for-age Z-scores
2. Incidence of stunting, underweight, and wasting

2. Hypotheses to be tested

1. Change in weight-for-age (WAZ), weight-for-length (WHZ), mid-upper arm circumference (MUAC)-for-age and head circumference-for-age Z-scores of infants provided with 10, 20, or 40 g/day of LNS from 6 to 18 months of age will be greater than that of infants who receive no dietary intervention at the same age.
2. Incidence of stunting, underweight, and wasting of infants provided with 10, 20, or 40 g/day of LNS from 6 to 18 months of age will be lower than that of infants who receive no dietary intervention at the same age.

3. Definition of the secondary growth outcomes

Change in weight-for-age (WAZ) and weight-for-length (WHZ) Z-scores rounded to two decimal points. Weight-for-age and weight-for-length at 6 (visit 0), 12 (Visit 26) and 18 months (visit 52) standardized by the WHO 2006 multi-centre growth standard. *(Form 00: Q2.5; Form 04: Q1.2, Q2.1, Q2.2, Q2.3, Q2.4)*

Change in length-for-age age Z-score (LAZ) rounded to two decimal points between 6 and 12 months of age (Figure 1). Length for age at 6 (visit 0) and 12 months (visit 26) standardized by the WHO 2006 multi-centre growth standard. *(Form 00: Q2.5; Form 04: Q1.2, Q2.2)*

Change in MUAC-for-age and head circumference-for-age Z-scores rounded to two decimal points at 6 (visit 0) and 18 months (visit 52) standardized by the WHO 2006 multi-centre growth standard. *(Form 00: Q2.5; Form 04: Q1.2, Q2.3, Q2.4)*

Incidence of stunting, underweight, and wasting. Moderate to severe stunting defined as LAZ <-2.0 and severe stunting defined as LAZ <-3.0 Z-score rounded to two decimal points. Incidence of stunting calculated at 12 and 18 months of age. Length for age at 12 months (visit 26) and 18 months (visit 52) standardized by the WHO 2006 multi-centre growth standard. *(Form 00: Q2.5; Form 04: Q1.2, Q2.2)*

Moderate to severe underweight defined as WAZ <-2.0 and severe underweight defined as WAZ <-3.0 Z-score rounded to two decimal points. Incidence of underweight calculated at 12 and 18 months of age. Weight for age at 12 months (visit 26) and18 months (visit 52) standardized by the WHO 2006 multi-centre growth standard. *(Form 00: Q2.5; Form 04: Q1.2, Q2.1)*

Moderate to severe wasting defined as WHZ <-2.0 and severe wasting defined as WHZ <-3.0 Z-score rounded to two decimal points. Incidence of wasting calculated at 12 and 18 months of age. Weight-for-length for age at 12 months (visit 26) and18 months (visit 52) standardized by the WHO 2006 multi-centre growth standard. *(Form 00: Q2.5; Form 04: Q1.2, Q2.1, Q2.2)*

4. Basis for the analysis: Intention to treat and per protocol

The basis for the analysis will be the same as that for the primary outcomes.

5. Time points for the analyses

All the above analyses will primarily cover the period from enrolment (when the participants are 6 months old) to the end of the intervention (when the participants are 18 months old).

6. Presentation of the study findings and hypothesis testing

6.1 Comparison of the anthropometric measurements at 18 months of age between intervention groups versus the control group

The group means and standard deviations for length (cm), weight (in kg), MUAC (cm), head circumference (cm) and LAZ, WAZ, WHZ, MUAC z-score and head circumference z-score (in z-score units) will be presented as indicated in Table 1. Group means and standard deviations for the change in anthropometric measurements and z-scores and their 95% confidence intervals between the control group and each of the intervention groups will be presented in Table 2.

Mean in anthropometrics, Z-score and change in measurements in each of the intervention groups will be individually compared against that of the control group using a two-sample t-test. The Holm's method will be used to adjust for the multiple comparisons (Aickin & Gensler,

1996). The raw P-values are presented in tables but the text and conclusion will be based on the Holm-adjusted P-values.

The null-hypothesis of LNS having no impact on linear growth is rejected for each comparison that yields a Holm's adjusted p-value <0.05.

6.3 Incidence of various forms of undernutrition at a single time point

The analysis will compare proportions demonstrating various forms of undernutrition at age 18 months (Table 3). Global null hypotheses of no differences between groups and pair-wise comparisons will be tested by Fisher's exact test. Pair-wise comparisons of intervention to control will be done if global null hypothesis is rejected with P<0.05. Participants who had the condition initially will be excluded from the analysis of that form of undernutrition.

7. General notes on statistical methods

7.1 Software

The same as that for the primary outcome analyses

7.2 Preparing anthropometric data for analysis

The same as that for the primary outcome analyses

7.3 Multiple comparisons

The same as that for the primary outcome analyses.

7.4 Confidence intervals

The same as that for the primary outcome analyses.

7.5 Interaction and effect modification

The same as that for the primary outcome analyses.

7.6 Covariate adjustment

The same adjustments will be done as for the main analyses.

8. References

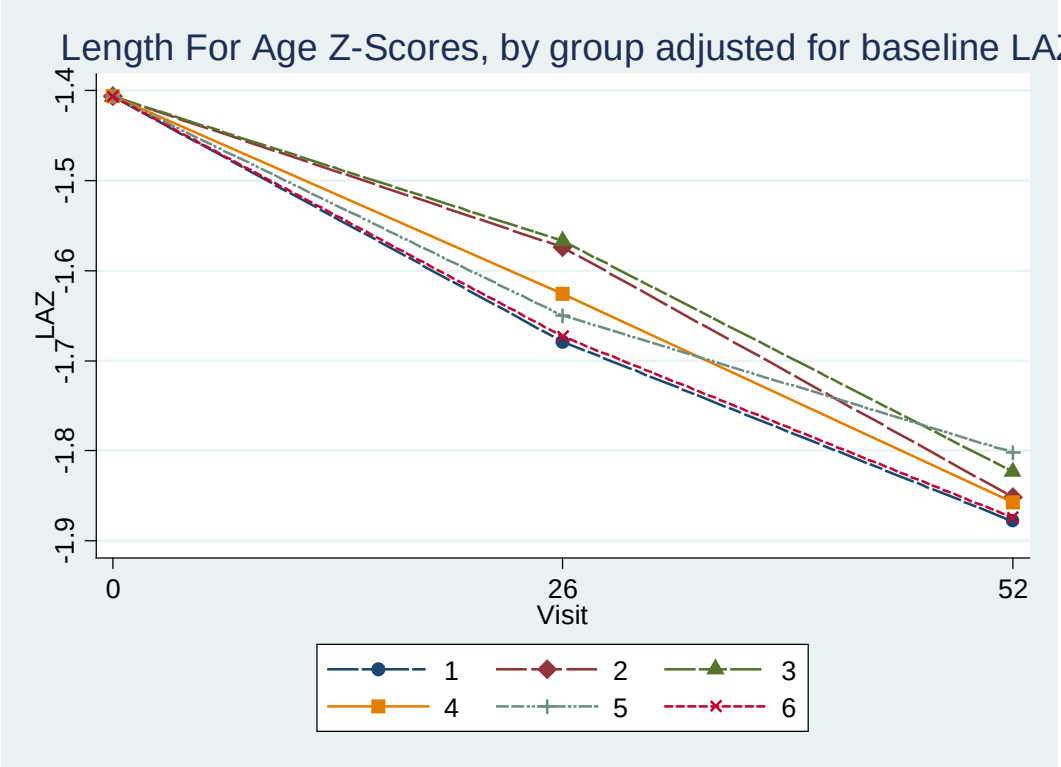
Aickin M, Gensler H. Adjusting for multiple testing when reporting research Results: The Bonferroni vs Holm Methods. *American Journal of Public Health* 1996; 86(5): 726-728.

9. Legends to the figures

Figure 1. Mean length-for-age z-score at 6, 12 and 18 months by intervention group

10. Figures

Figure 1. Mean length-for-age z-score at 6, 12 and 18 months by intervention group



11. Tables

Table 1 Comparison in anthropometric in the intervention groups at 18 months

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**
Mean weight (SD)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean length (SD)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean middle upper arm circumference (SD)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean head circumference (SD)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean Weight-for-age z-score (SD)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)

Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean Length-for-age z-score (SD)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean Weight-for-length z-score (SD)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean MUAC z-score (SD)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Mean head circumference z-score (SD)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Table 2 Anthropometric outcome changes during and up to 12 months of intervention

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**
Change in weight, mean (SD),kg	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in length, mean (SD), cm	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in Middle upper arm circumference, mean (SD),cm	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in Head circumference, mean (SD), cm	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in Weight-for-age z-score, mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in Length-for-age z-score mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in head circumference z-score mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in MUAC z-score mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)
Change in Weight-for-length z-score mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx)
Raw p-value for intervention compared to control		x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Difference in means (95% CI)		xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)

Table 3 Incidence of various forms of malnutrition at 18 months of age

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**P-value**
Proportion severe stunting (LAZ<-3), No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	x.xxx
Proportion moderate to severe stunting (LAZ <-2), No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	x.xxx
Proportion severe underweight (WAZ <-3), No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	x.xxx
Proportion moderate to severe underweight (WAZ <-2), No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	x.xxx
Proportion severe wasting (WHZ <-3), No. /total No. No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	xxx/xxx (xx.x%), x.xx (x.xx-x.xx)	x.xxx
Proportion moderate to severe wasting (WHZ <-	xxx/xxx	xxx/xxx (xx.x%),	xxx/xxx (xx.x%),	xxx/xxx (xx.x%), x.xx	xxx/xxx (xx.x%),	xxx/xxx (xx.x%),	x.xxx

2), No. /total No. (%) Relative risk (95% CI)	(xx.x%)	x.xx (x.xx-x.xx)	x.xx (x.xx-x.xx)	(x.xx-x.xx)	x.xx (x.xx-x.xx)	x.xx (x.xx-x.xx)	

Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Appendix 02: Developmental outcomes at age 18 months (added on 07.05.2013)

Contents

1. Study objectives

The main aim of the trial was to assess the safety and impact of supplementation with Lipid-based Nutrient Supplements (LNS) on linear growth of infants and young children and to identify individual, household, and village-level characteristics that modify the effects of LNS on child growth. A secondary aim is to similarly study the impact of LNS on various other (secondary) outcomes in the same target group.

The aim of the secondary analyses described in this appendix is to compare infants in 6 different intervention groups: dietary supplementation from age 6 to 18 months with 10 g, 20 g or 40 g per day milk-containing LNS, or 20 g or 40 g per day milk-free LNS, or nothing (delayed intervention) on the following outcomes:

1. 18-month motor development, language development, socio-emotional development, executive function, and interaction with caregivers
2. Prevalence of severe and moderate to severe delay in motor development, language development, socio-emotional development, and executive function

2. Hypotheses to be tested

1. 18-month scores in gross and fine motor development, language development, socio-emotional development, executive function, and interaction with caregivers of infants provided with 10, 20, or 40 g/day of LNS from 6 to 18 months of age will be greater than that of infants who receive no dietary intervention at the same age.

2. 18-month scores in gross and fine motor development, language development, socio-emotional development, executive function, and interaction with caregivers of infants receiving 20 or 40 g/day of LNS without milk from 6 to 18 months of age will not be lower than that of infants receiving a comparable intervention with milk-containing LNS.

3. 18-month scores in gross and fine motor development, language development, socio-emotional development, executive function, and interaction with caregivers of infants provided with 20 or 40 g/day of LNS from 6 to 18 months of age will be greater than that of infants who receive 10g/day and that of infants provided with 40 g/day will be greater than that of infant who receive 20 g/day.

4. Hypotheses 1-3 will also be examined with regard to the prevalence of severe and moderate to severe delay in motor development, language development, socio-emotional development, and executive function.

3. Definition of the 18-month developmental outcomes

The gross motor score is calculated as the sum of 35 Kilifi Developmental Inventory (KDI) gross motor items, each scored 0 or 1 (sum of *Form 25a Q 5.1-5.2* and *5.6-7.12*). Severe delay is defined as the bottom 10% of our sample. Moderate to severe delay is defined as the bottom 25% of our sample.

The fine motor score is calculated as the sum of 34 KDI fine motor items, each scored 0 or 1 following Abubakar et al. (2008). Severe delay is defined as the bottom 10% of our sample. Moderate to severe delay is defined as the bottom 25% of our sample.

The psychomotor score is calculated as the sum of 69 KDI fine and gross motor items, each scored 0 or 1. Severe delay is defined in two ways: (1) the bottom 10% of our sample and (2) <-3 SD below the mean according to published norms from Kenya (Abubakar et al. 2008). Moderate to severe delay is defined in two ways: (1) the bottom 25% of our sample and (2) <-2 SD below the mean according to published norms from Kenya (Abubakar et al. 2008).

Language development is quantified as

a. Vocabulary score, calculated as the sum of *Form 25c LANGVOCAB1* through *LANGVOCAB100.* Severe delay is defined as the bottom 10% of our sample. Moderate to severe delay is defined as the bottom 25% of our sample.
b. Expressive vocabulary > 10 words vs. <= 10 words, derived from the vocabulary score
c. Word combining (Has the child started combining words into sentences? 0 = not yet, 1 = sometimes, 2 = often) *Form 25c Q 4.1*

Socio-emotional development is calculated as the sum of *Form 25b PSED1* through *PSED19.* Severe delay is defined as the top 10% of our sample (a lower score indicates more advanced socio-emotional development). Moderate to severe delay is defined as the top 25% of our sample.

Executive function is calculated as

a. A not B task total number correct, *Form 25a Q16.2.* Severe delay is defined as the bottom 10% of our sample. Moderate to severe delay is defined as the bottom 25% of our sample.
b. A not B task total errors after set 1, *Form 25a Q16.3*
c. A not B task total trials completed, *Form 25a Q 16.1*. If this variable is not normally distributed, another statistical approach will be used, such as creating a dichotomous variable

Interaction with caregivers is calculated as the sum of the activities with adults in the past three days (*Form 25d Q 4.1.1* through *Q 4.5.3*).

4. Basis for the analysis: Intention to treat and per protocol

The basis for the analysis will be the same as that for the primary outcomes.

5. Presentation of the study findings and hypothesis testing

5.1 Comparison of the developmental scores at 18 months of age between intervention groups versus the control group

The group means and standard deviations for the gross motor score, fine motor score, psychomotor score, vocabulary score, socio-emotional score, and A not B task total number correct, total errors after set 1, and total trials completed, and the interaction with caregivers score will be presented as indicated in Table 1. Both raw scores and z-scores will be presented. The results of pairwise comparisons will be indicated by superscripts. Means that are significantly different from each other will be marked by different letters (e.g., a and b). Means that are not significantly different from each other will be marked by the same letter.

We will use a one-factor analysis of variance (ANOVA) (or ANCOVA in analyses that include covariates, as specified in section 6.6) to test for differences between the six groups. If the group effect is significant at the level of $p < .07$, we will use Tukey-Kramer's test for post-hoc pairwise comparisons between each intervention group and the control group. The null-hypothesis of LNS having no impact on development will be rejected for each comparison that yields a Tukey-Kramer's adjusted p-value < 0.05.

5.2 Comparison of milk-containing and milk-free LNS

We will use a two-factor ANOVA/ANCOVA with main effects of dose of LNS (0g, 10g, 20g, or 40g) and milk (yes or no) and the dose by milk interaction. The control and 10g LNS groups will be treated as containing milk. This means that the main effect of milk and the dose by milk interaction are calculated in the model using the data from the 20g and 40g groups only.

If the dose by milk interaction is not significant at the level of $p < 0.1$, we will conclude that the effect of milk does not differ between the 20g and 40g groups. In this case, the difference in mean scores between the milk-containing and milk-free LNS will be presented as shown in Table 2. A 90% CI will be created for the milk effect, which will be the average of the individual milk effects for the 20g and 40g groups. The set non-inferiority margin for this analysis will be 0.2 Z-score units. If the lower bound of the 90% CI of the coefficient is larger than -0.2, non-inferiority is confirmed. If the upper bound of the 90% CI is smaller than -0.2, inferiority is confirmed. If

the upper and lower bounds of the 90% CI are larger and smaller, respectively, than -0.2, the finding is not conclusive.

If the dose by milk interaction is significant at the level of $p < 0.1$, we will examine the effect of milk separately in the 20g and 40g groups. In this case, the difference in mean scores between the milk-containing and milk-free LNS will be presented in an expanded version of Table 2 which will present the coefficient, 90% confidence interval, and p-value for (1) the difference in scores in 20g/day LNS with milk versus 20g/day LNS without milk, (2) the difference in scores in 40g/day LNS with milk versus 40 g/day LNS without milk 40g, and (3) the dose by milk interaction term. In this case, a 90% CI will be created for each of the individual milk effects for the 20g and 40g groups. The set non-inferiority margin for this analysis will be 0.2 Z-score units. If the lower bound of the 90% CI of the coefficient is larger than -0.2, non-inferiority is confirmed. If the upper bound of the 90% CI is smaller than -0.2, inferiority is confirmed. If the upper and lower bounds of the 90% CI are larger and smaller, respectively, than -0.2, the finding is not conclusive.

5.3 Comparison of the 10g, 20g, and 40g LNS groups

If the dose by milk interaction is significant at the level of $p < 0.1$, we will use Tukey-Kramer's test for post-hoc pairwise comparisons between each intervention group in the one-factor ANOVA/ANCOVA described in section 5.1. The null-hypothesis of higher doses of LNS having no impact compared to lower doses will be rejected for each comparison that yields a Tukey-Kramer's adjusted p-value < 0.05.

If the dose by milk interaction is not significant at the level of $p < 0.1$ and if we conclude that LNS without milk is not inferior to milk-containing LNS, we will use Scheffe test in the one-factor ANOVA/ANCOVA described in section 5.1 to compare the mean scores among the combined 40g group (with and without milk), the combined 20g group (with and without milk), the 10g group, and the control group. The null-hypothesis of higher doses of LNS having no impact compared to lower doses will be rejected for each comparison that yields a Scheffe's adjusted p-value < 0.05.

If the dose by milk interaction is not significant at the level of $p < 0.1$ and if we do not conclude that LNS without milk is not inferior to milk-containing LNS, we will use Tukey-Kramer's test for post-hoc pairwise comparisons between each intervention group in the two-factor ANOVA/ANCOVA described in section 5.2. This model calculates the effect of dose while controlling for the effect of milk. The null-hypothesis of higher doses of LNS having no impact compared to lower doses will be rejected for each comparison that yields a Tukey-Kramer's adjusted p-value < 0.05.

5.4 Prevalence of severe and moderate to severe developmental delay

The proportions of children demonstrating severe and moderate to severe developmental delays at age 18 months will be presented as shown in Table 3. We will use logistic regression,

following the same approach outlined in sections 5.1 through 5.3. The results of pairwise comparisons will be indicated by superscripts. Relative risks and confidence intervals will be calculated based on the method described in Kleinman (2009).

6. General notes on statistical methods

6.1 Software

SAS for Windows Release 9.3 (Cary, NC) will be used for all analyses.

6.2 Calculating scores and z-scores

If a large percentage of data is missing for any item, we will exclude that item from the total score. For all other missing item scores, we will impute the scores based on the other items in the same subscale. We will use the imputation method described in Raghunathan et al. (2001).

Z-scores of developmental variables will be calculated based on the distribution of the iLiNS-DOSE sample, by standardizing the distribution to a mean of 0 and standard deviation of 1.

6.3 Multiple comparisons

The Tukey-Kramer adjustment method is used.

6.4 Confidence intervals

The same as that for the primary outcome analyses.

6.5 Interaction and effect modification

We will examine the same factors as that for the primary outcome analyses. In addition, we will examine the following effect modifiers:

1. Child stunting at enrollment (LAZ < -2)
2. Child wasting at enrollment (WHZ <-2)
3. Child 6 month iron deficient anemic, iron deficient non-anemic, non-iron deficient anemic, non-iron deficient non-anemic
4. Maternal education (we will examine the data to establish a cut-off)
5. Family care indicators (we will examine the data to establish a cut-off)

6.6 Covariate adjustment

The same procedure will be followed as for the main analyses, with three models rather than four models:

1. No covariate adjustment

2. Adjustment for child age at developmental assessment
3. Adjustment for child age at developmental assessment, baseline LAZ-score, and baseline WHZ-score and for any of the variables presented in tables 1a and 1b of the primary outcome Statistical Analysis Plan (SAP) showing statistically significant association (at $p<0.1$ level) with the developmental score

In addition to the variables in tables 1a and 1b of the primary outcome SAP, we will consider the following variables for inclusion:

1. Family care indicators score, if this score is not different between supplement groups.
2. For the KDI scores, the child's mood, interaction with the tester, and activity level during testing, if they are not different between supplement groups (*Form 25a Q 3.1 - 3.3*).
3. For the language scores, the child's primary language (Chichewa, Chiyao, English, or other) and the number of languages to which the child had been exposed (*Form 25c Q 1.7-1.8*).

7. References

Abubakar, A., Holding, P. A., Van Baar, A., Newton, C. R. J. C., & Van de Vijver, F. J. R. (2008). Monitoring psychomotor development in a resource-limited setting: An evaluation of the Kilifi Developmental Inventory. *Annals of Tropical Pediatrics*, 28, 217-226.

Kleinman, L. C., & Norton, E. C. (2009). What's the Risk? A simple approach for estimating adjusted risk measures from nonlinear models including logistic regression. *Health services research, 44*(1), 288-302.

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology, 27(1), 85-95.

8. Tables

Table 1 Results of the ANCOVA comparing developmental scores in the intervention groups at age 18 months.

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**P-value**
Mean gross motor raw score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean gross motor z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean fine motor raw score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean fine motor z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean psychomotor raw score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean psychomotor z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean vocabulary raw score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean vocabulary z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean socio-emotional raw score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean socio-emotional z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean A not B task total correct (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean A not B task total correct z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	

Mean A not B task errors (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean A not B task errors z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean A not B task total trials completed (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean A not B task total trials completed z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	
Mean interaction with caregivers raw score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	x.xx
Mean interaction with caregivers z-score (SD) [y]	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	xx.x[x] (xx.x)	

[y]Footnotes will report the covariates included for each outcome.

Table 2. Regression results of the difference between the milk-containing and milk-free LNS

Variable	**Coefficient**	**90 % CI**	**P-value**
Difference in mean gross motor z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean fine motor z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean psychomotor z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean vocabulary z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean socio-emotional z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean A not B task total correct z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean A not B task errors z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean A not B task total trials z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx
Difference in mean interaction with caregivers z-score in milk-free LNS compared to milk-containing LNS	xx.xx	xx.xx to xx.xx	x.xx

Table 3 Results of the logistic regression comparing the prevalence of developmental delay in the intervention groups at 18 months of age.

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**P-value**
Proportion severe gross motor delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion moderate to severe gross motor delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion severe fine motor delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion moderate to severe fine motor delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion severe psychomotor delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx

Proportion moderate to severe psychomotor delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion severe language delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion moderate to severe language delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion severe socio-emotional delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion moderate to severe socio-emotional delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx
Proportion severe executive function delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx

Proportion moderate to severe executive function delay, No. /total No. (%) Relative risk (95% CI)	xxx/xxx (xx.x%)[x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	xxx/xxx (xx.x%), x.xx (x.xx-x.xx) [x]	x.xx

Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Appendix 03: Breast milk intakes (added on 07.05.2013)

Contents

Study objective

To estimate breast milk intake in a sub-sample of Malawian infants participating in the LNS intervention trial at 9-12 months of age using the dose-to-mother deuterium oxide dilution technique.

1. Hypothesis to be tested

1. Mean breast milk intake (g/d) of 9-12 rural Malawian infants supplemented with 10-40 g/d of LNS is not lower than that of infants not supplemented with LNS.

2. Definition of the outcomes

Breast milk intake will be measured from deuterium oxide enrichment data using the solver function in excel. The breast milk intake estimate for each infant is calculated based on the two compartment steady state model between the mother and the infant (Shipley and Clark, 1972 as reported in IAEA manual for breast milk intake assessment). The output obtained after running the solver function is the mean breast milk intake (g/day) over a 14 day period. Breast milk intake will also be expressed as grams per kilogram body weight.

Mean non-breast milk oral water intakes and total daily water intake will be estimated in the same way as mean breast milk intakes.

3. Basis for the analysis: Intention to treat and per protocol

Primarily, the analysis will be based on the principle of modified intention-to-treat. The modification concerns six participants who were accidentally allocated to another group than originally randomized, as explained in the main analysis plan. Another modification to the intention to treat principle is that individuals with missing and incomplete data (as defined by participants from whom not all planned 7 time point saliva samples and weight measurements were taken) will be removed from the analysis.

As supplementary evidence, we will perform a per protocol analysis. For this analysis, participants will not be grouped based on their intended intervention, but rather by their observed actual LNS intakes. The actual intakes were determined with a modified 24 h recall on two separate days (approximately one week apart) during the breast milk intake assessment. The assessment categorized participants into four groups based on their actual LNS intake on the study days: None, 1-10 g, 11-20 g, and >20 g. In the per protocol analysis, we will use these

groups for the comparisons. Since not all the infants in breast milk intake study have dietary data, the per protocol analysis will be limited to participants with both breast milk and dietary intake data.

4. Time points for the analyses

The analysis for the study will primarily cover a two week period from when the participating infants are 9 to 12 months old.

5. Presentation of the study findings and hypothesis testing

6.1 Comparison of breast milk intakes and non-breast milk oral water intakes, among the intervention groups at 9-12 months

Box-whisker plots of the mean breast milk intakes in the 4 groups will be shown side by side in Figure 2. The group means and standard deviations for the intakes of breast milk, non-breast milk water and total water (in grams / day) will be presented as indicated in Table 2. Table 3 will tabulate the difference in mean breast milk intakes and their 90% confidence intervals between the control group and each of the intervention groups. Ninety per cent instead of 95% CI will be used because the non-inferiority consideration is one-sided and the use of 90% CI is the convention in this situation (Senn, 1997).

Evaluation

The set non-inferiority margin for this analysis will be 10% of the daily energy needs of infants corresponding to breast milk intake among 9-12 mo old infants. If the upper bound of the 90% CI for the point estimate for the difference in means (control mean minus intervention group mean) is smaller than the set margin of 10%, non-inferiority is confirmed. For instance if the difference between the control and any of the intervention is 40g with confidence interval of 30g to 60g, then that particular LNS dose is non-inferior to the control since the upper bound, 60g, is equivalent to about 6 % which is less that the non-inferiority margin (10%). If the lower bound of the 90% CI is larger than 10%, inferiority is confirmed. For example if the CI of the difference between the control and any of the LNS dose is 120g to 130g, since 120g is equivalent to about 11.8% which is above the noninferiority margin, the LNS dose in question would be considered inferior. If the upper and lower bounds of the 90% CI are larger and smaller, respectively, than 10%, the finding will be considered inconclusive for target group inference.

The set non-inferiority margin used in the present study (10% of the breast milk intake among control children) is based on clinical judgment and represents less than 10% of total daily energy needs of infants at 9-11 mo of age (Dewey and Brown 2003).

6. General notes on statistical methods

6.1 Software

Analyses will be done in Stata version 12 and SAS for Windows version 9.3 (Cary, NC).

6.2 Preparing breast milk data for analysis

The first procedure in obtaining the breast milk intakes and non-breast milk oral water intake is to transfer the data from the solver output for each participant onto an excel file. Thus for all the participants their data were transferred to a common excel sheet from which suspicious values would be identified. Calculating the means for breast milk, non-breast milk oral intake and total water intake requires complete deuterium enrichment data to be used in the two compartment steady state model solver function in excel, this implies there are no missing data for participants who completed the study protocol.

6.3 Multiple comparisons

The initial plan is to compare means for intakes of breast milk, non-breast milk oral water and total water intakes using the global null hypothesis of no difference between groups. A trend analysis will then be conducted to find out if those receiving the dietary supplements at 10-40 g/d differ from the control group, which did not receive the study dietary supplements.

6.4 Confidence intervals

The confidence intervals (CI) at 95% level will be provided for all the three main outcomes. The general group level comparison will also contain 95% CI which will be used to compare against the reported global mean intakes of breast milk, non-breast milk oral water intake and total water intake. Breast milk intake in the study will be considered significantly different if the lower CI of the group intakes is higher than the upper CI for the global reported intake. Confidence intervals will be based on t-test.

6.5 Covariate adjustment

Breast milk intakes, non-breast milk oral water intake, total water intake will be adjusted for infant baseline anthropometric indices.

7. References

Dewey, K. G. & Brown, K. H. Update on technical issues concerning complementary feeding of young children in developing countries and implications for intervention programs. *Food and nutrition bulletin* **24**, 5-28 (2003).

Senn S. *Statistical Issues in Drug Development*. Chichester, GBR. John Wiley & Sons. 1997.

Shipley, R.A., Clark, R.E., Tracer Methods for in Vivo Kinetics. Theory and Applications, Academic Press, New York and London (1972).

8. Legends to the figures

Figure 1. Participant flow

Figure 2. Box-Whisker plots of breast milk intakes by groups

9. Figures

Figure 1: Participant flow

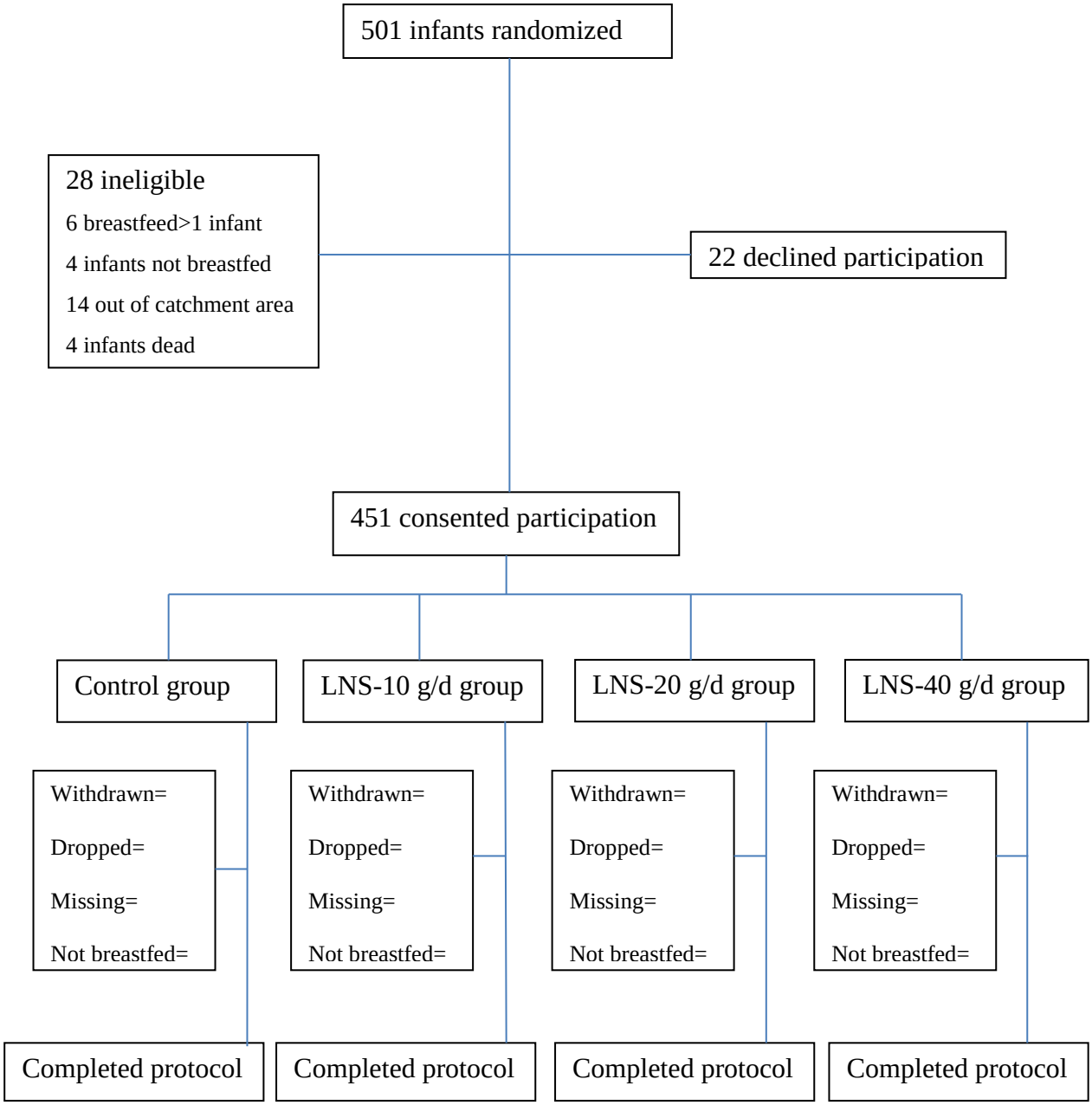
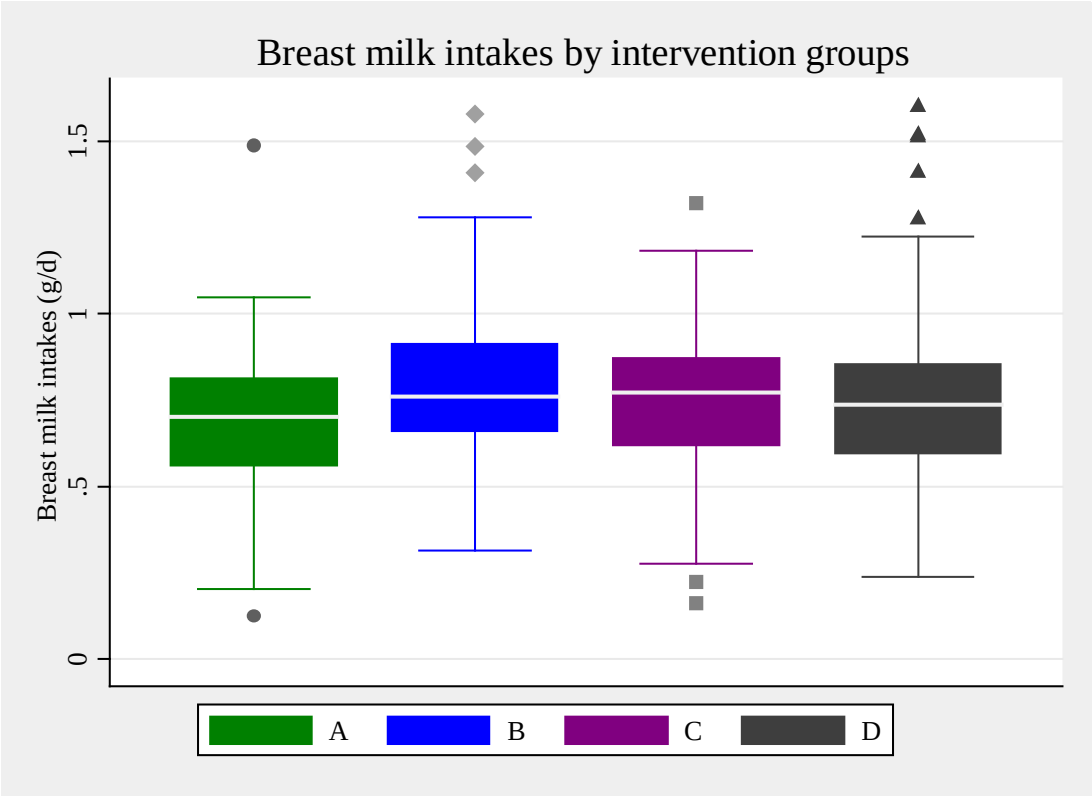


Figure 2. Box-Whisker plots of breast milk intakes by the intervention group



10. Tables

Table 1 Comparison in baseline characteristics

Variable	**Control**	**10g milk LNS**	**20g LNS**	**40g LNS**	**Test**
Number of participants	xxx	xxx	xxx	xxx	
Infants characteristics					
Age, months (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Proportion of Males (percentage)	xxx / xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	xxx/ xxx (xx%)	Chi-squared
Weight, kg (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Maternal characteristics					
Age, year (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Height, kg (mean, SD, N)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	ANOVA
Weight, kg (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
BMI, kg/m^2 (mean, SD, N)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA

Table 2 Comparison of breast milk intakes between children in the control and intervention groups, intention-to-treat analysis

	Planned daily dose of LNS, grams / day			
Variable	**Control, 0g LNS**	**10g LNS**	**20g LNS**	**40g LNS**
Mean (SD) breast milk intake, grams / day	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)
Difference (95% CI) in mean intakes between the indicated intervention group andthe control		x.xx	x.xx	x.xx
Mean non-breast milk oral water intake (SD)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)
Difference (95% CI) in mean intakes between the intervention and control groups		x.xx	x.xx	x.xx
Mean total water intake (SD)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)
Difference (95% CI) in mean intakes between the intervention and control groups		x.xx	x.xx	x.xx

Table 3 Comparison of breast milk intake between children consuming different daily doses of LNS, per protocol analysis

	Actual daily intake of LNS, grams / day			
Variable	**0**	**1 – 10**	**11 – 20**	**> 20**
Mean (SD) breast milk intake, grams / day	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)
Difference (95% CI) in mean intakes between the indicated intervention group and the control		x.xx	x.xx	x.xx
Mean non-breast milk oral water intake (SD)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)
Difference (95% CI) in mean intakes between the intervention and control groups		x.xx	x.xx	x.xx
Mean total water intake (SD)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)	xxx.x (xx.x)
Difference (95% CI) in mean intakes between the intervention and control groups		x.xx	x.xx	x.xx

Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Appendix 04: The effect of LNS on physical activity (added on 24.05.2013)

Table of Contents

1 Study objectives

The main aim of sub-study is to assess impact of LNS supplementation on physical activity of children at the age of 18 months. This will be done by comparing physical activity of children in the six intervention groups: dietary supplementation with 10 g, 20 g or 40 g/day milk-containing LNS, 20 g or 40 g/day milk-free LNS or nothing (delayed intervention).

A secondary aim is to assess dose response between LNS supplementation and physical activity. In addition, we will explore possible effect modifiers of the impact of LNS supplementation on physical activity.

2 Hypotheses to be tested

The mean physical activity of infants provided with 10, 20, or 40 g/day of LNS from 6 to 18 months of age is greater than that of infants who receive no dietary intervention at the same age.

3 Definition of the physical activity outcomes

Primary outcome: mean accelerometer counts

Physical activity counts used in the analysis are vector magnitude counts, calculated by taking the square root of the sum of squared activity counts of each three axis. The mean counts/15 s of each day will be averaged over all valid days to produce mean of means for each participant.

Data for physical activity will be considered missing if the actual measurement date was over 30 days from the target date.

Secondary outcomes

Percentage of time spend in moderate-to-vigorous physical activity (MVPA) is averaged over all valid days and the averaged value (per participant) is used in the analysis. MVPA is defined in two ways: 1) with vector magnitude counts of ≥ 208/15 s (Pulakka et al, unpublished) and 2) with vertical axis activity counts ≥ 419 counts/15 s (Trost et al. 2011). Pulakka vector magnitude cut point is used because it is the only cut point developed for vector magnitude for children this age. Trost cut point for vertical axis is used to allow comparison with previous studies using that cut point and older models of accelerometers with only vertical axis readings.

For mean vertical axis accelerometer counts/15 s, mean counts of each day are averaged over all valid days and the average value is used in the analyses.

4 Basis for the analysis: Intention to treat and per protocol

The basis for the analysis will be the same as for the primary outcomes, i.e. intention-to-treat analysis for all except for the six participants who were accidentally allocated to another group than actually randomized. Those six participants will be analyzed in the group that they were erroneously allocated to.

5 Time points for the analyses

All the above analyses will be done at the end of the intervention (when the participants are 18 months old).

6 Presentation of the study findings and hypothesis testing

6.1 Success of enrolment, follow-up and physical activity measurement

All registered participants and the success of their follow-up, including physical activity measurement, will be described in a flow chart (Figure 1). For additional information, drop-out rate (including participants for whom enough accelerometer data was not available) between groups will be tested with Fisher's exact test and baseline characteristics of drop-outs compared to those who completed the study will be tested with t-test or Fisher's exact test. P-values for these tests will be shown in the text.

6.2 Baseline information

Participant characteristics at the trial enrollment (age of 6 months) and at physical activity measurement (at 18 months) will be tabulated by treatment arms as indicated in Table 1. Hypothesis testing will be performed for baseline information to give additional information but p-values will not be presented in Table 1 of the eventual manuscript. Methods used for hypothesis testing are indicated in Table 1.

6.3 Comparison of physical activity between intervention groups versus the control group

Box-whisker plots of the mean vector magnitude accelerometer counts/15s of all 6 groups will be shown side by side in Figure 2. The group means and standard deviations for the mean counts and % of time spend in MVPA for vector magnitude and vertical axis will be presented as indicated in Table 2. The table will also tabulate the difference in activity indicators (mean vector magnitude counts, mean vertical activity counts, % time in MVPA) and their 95% confidence intervals between the control group and each of the intervention groups.

The activity indicators in each of the intervention groups will be individually compared against that of the control group using two-sample t-test. The raw P-values are presented in the tables and text. In addition, the P-values adjusted to multiple comparisons with Holm's method will be presented in the text for the main outcome, mean vector magnitude accelerometer counts/15 s.

7 General notes on statistical methods

7.1 Software

All analyses will be done in Stata/IC version 11. The WHO 2006 multi-centre growth standard will be used for age-and-sex standardization of weight, length (height), weight-for-height, MUAC and head circumference.

7.2 Preparing physical activity data for analysis

Data that was originally compiled by ActiLife software (version 5) from ActiGraph GT3X+ devices, will be extracted and combined using the following procedure:

- .gt3x files are converted to .agd files (with 3 axes and 15s epoch length) in ActiLife software
- .agd files are converted to .csv files in ActiLife software
- .csv files brought to StataIC software (version 11) and compiled to the same .dat file.
- Strings of consecutive zeroes of 20 minutes or more are deleted as well as night time (between 8:00 p.m. and 5:00 a.m.)
- First and last day are also deleted as incomplete days

The data is used for the analyses if the participant has minimum of 4 days of minimum of 6 hours of data.

7.3 Preparing anthropometric data for analysis

The same as for the primary outcome analysis

7.4 Multiple comparisons

The same as for the primary outcome analysis

7.5 Confidence intervals

The same as for the primary outcome analysis

7.6 Interaction and effect modification

The following variables will be tested for interaction between the intervention group and the primary outcome (mean vector magnitude accelerometer counts). All tests will be done using the likelihood ratio test. The variables tested could logically modify the effect of the nutritional intervention on infancy and physical activity. Variables included (as continuous variables where possible) in this analysis include:

1. The participant's baseline length-for-age (below / above sample median) at 6 months
2. The participant's baseline weight-for-length (below/above sample median) at 6 months
3. The participant's sex
4. Season of participant's birth
5. Maternal education

6. Maternal age
7. Household food security

If a statistically significant interaction ($p<0.1$) is found, the outcome analysis will be completed as stratified by the respective predictor variable.

7.7 Covariate adjustment

The main analysis, the results of which will be shown in tables and figures, will be completed without any covariate adjustments.

As a secondary analysis we will construct a regression model for physical activity, adjusting for the participant's LAZ and WHZ at 6 enrolment, sex, season of enrolment, maternal education, maternal age, and household food insecurity. However, if any of these variables is found to be an effect modifier (see chapter 7.6), it will not be included in the model.

As a sensitivity test for the latter analysis, we will use two alternative methods to build the regression model:

1. Inclusion in the model of also the effect modifiers and respective interaction terms.
2. Inclusion in the model of only those variables that are associated with physical activity (mean vector magnitude accelerometer counts) at $p<0.1$ level) with.

8 References

Maldonado G, Greenland S. Simulation Study of Confounder-Selection Strategies. *American Journal of Epidemiology* 1993; 138(11): 923-936.

Trost SG, Fees SF, Haar SJ, Murray AD, Crowe LK. Identification and Validity of Accelerometer Cut-Points for Toddlers. *Obesity* 2012; 20(11): 2317-2319

9 Legends to the figures

Figure 1. Participant flow

Figure 2. Box-Whisker plots of time in moderate-to-vigorous physical activity by groups

10 Figures

Figure 1. Participant flow

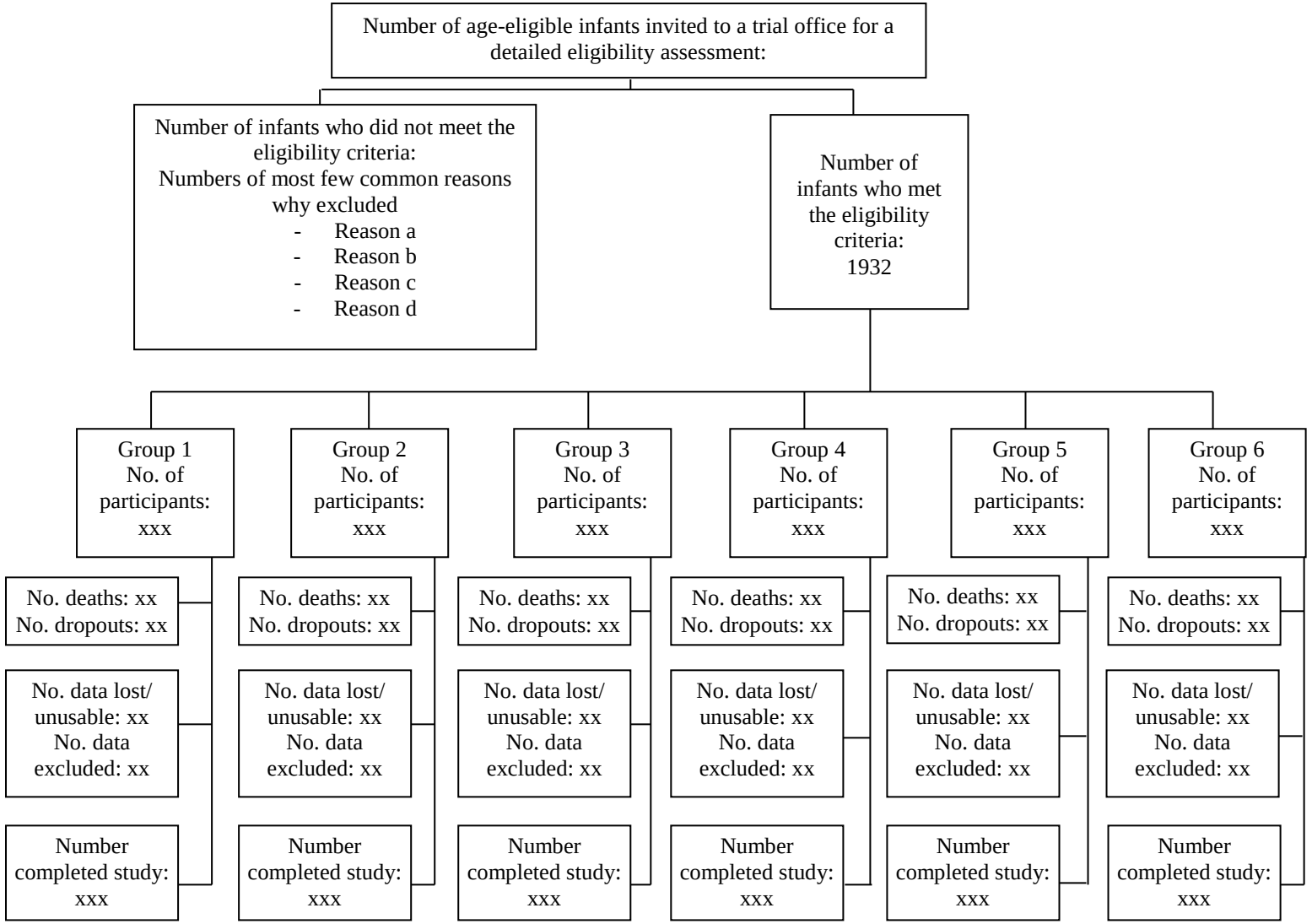
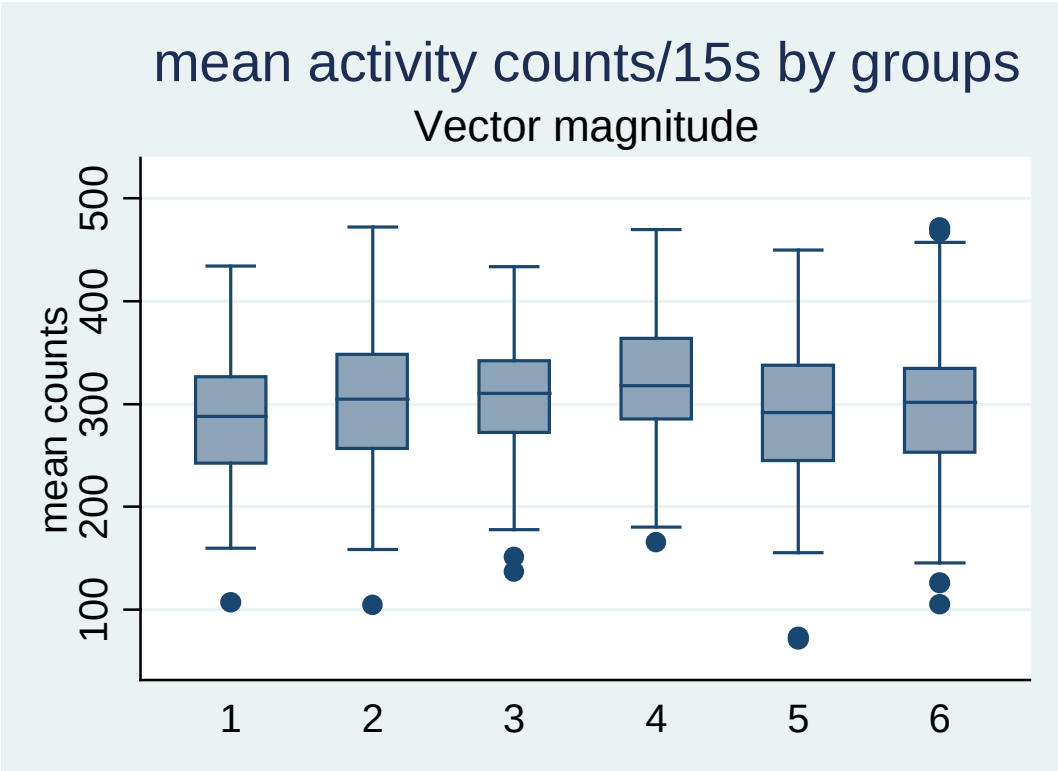


Figure 2. Box-Whisker plots of time in moderate-to-vigorous physical activity by groups



Tables

Table 1 Baseline characteristics of participants at baseline and 18 months

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**Test**
Number of participants	xxx	xxx	xxx	xxx	xxx	xxx	
Mean (SD) maternal age at trial enrollment, years	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Mean (SD) maternal education at trial enrollment, completed years at school	x.x (x.x)	x.x (x.x)	x.x (x.x)	x.x (x.x)	x.x (x.x)	x.x (x.x)	ANOVA
Proportion of food insecure households at trial enrollment	xx.x%	xx.x%	xx.x%	xx.x%	xx.x%	xx.x%	Fisher's exact test
Percentage of males	xx.x%	xx.x%	xx.x%	xx.x%	xx.x%	xx.x%	Fisher's exact test
Season of child's birth	I: xx.x% II: xx.x% III: xx.x% IV: xx.x%	I: xx.x% II: xx.x% III: xx.x% IV: xx.x%	I: xx.x% II: xx.x% III: xx.x% IV: xx.x%	I: xx.x% II: xx.x% III: xx.x% IV: xx.x%	I: xx.x% II: xx.x% III: xx.x% IV: xx.x%	I: xx.x% II: xx.x% III: xx.x% IV: xx.x%	Chi-squared
Mean (SD) age at physical activity measurement, months	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	ANOVA
Mean (SD) length-for-age z-score at trial enrollment	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Mean (SD) length-for-age z-score at physical activity measurement	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Mean (SD) weight-for-length z-score at trial enrollment	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA
Mean (SD) weight-for-length z-score at physical activity measurement	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	xx.xx (xx.xx)	ANOVA

Table 2 Physical activity at the trial groups

Variable	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**
Mean (SD) vector magnitude accelerometer counts/ 15 s	xxx (xx)	xxx (xx)	xxx (xx)	xxx (xx)	xxx (xx)	xxx (xx)
Difference (95% CI) in mean vector magnitude accelerometer counts between the indicated intervention and the control group.		xx.x (xx.x to xx.x), p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx
Mean (SD) vertical axis accelerometer counts/15 s	xxx (xx)	xxx (xx)	xxx (xx)	xxx (xx)	xxx (xx)	xxx (xx)
Difference (95% CI) in mean vertical axis accelerometer counts between the indicated intervention and the control group.		xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx
Mean (SD) % of time in MVPA, vector magnitude	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)
Difference (95% CI) in % of time in MVPA, vector magnitude, between the indicated intervention and the control group.		xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx
Mean (SD) % of time in moderate-to-vigorous physical activity, vertical axis	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)
Difference (95% CI) in % of time in MVPA, vertical axis, between the indicated intervention and the control group.		xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx	xx.x (xx.x to xx.x) p=x.xxx

MVPA, moderate-to-vigorous physical activity

Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Appendix 05: Effect on infant and young child feeding practices (version 1.0 added on 26.07.2013, revised to version 2.0 on 28.04.2014)

Contents

1. Version history

Version number	Version date	Prepared by	Description of the completed editions
01.0	26.07.2013	Arimond	Original document Appendix 05 added
01.1	16.04.2014	Arimond	Updated Appendix 05 to delete detailed power calculations and to add equivalence testing for quasi-continuous outcomes. Several minor (non-substantive) corrections were made to description of variable construction. Tables are edited to defer decision on presentation of medians vs. means until by-group distributions can be examined.

2. Overview and study objectives

The analysis presented here is nested within a pre-existing iLiNS-DOSE analysis plan for primary and other secondary outcomes. Refer to the main analysis plan for: inclusion and exclusion criteria for the trial; data cleaning protocols; procedures for breaking code; and procedures for modifying this protocol.

The main objective of data collection related to IYCF practices is to compare practices across intervention groups. This analysis is motivated by concerns that energy-dense LNS may displace breastfeeding and/or nutrient-dense local foods and/or impede dietary diversification with local foods, thus negatively impacting infant feeding practices and development of infant dietary preferences and habits. Effects on IYCF practices could be mediated either by maternal perceptions of different needs for breast milk or local foods for infants receiving supplements and/or by a change in appetite, demand for breastfeeding, or preference for local foods among infants who consume the supplement.

Data on IYCF practices have been gathered on the full study samples, and complement quantitative dietary data and breast milk intake data, which were collected on sub-samples and at fewer time points.

IYCF practices we will compare across groups include: continued breastfeeding, frequency of breastfeeding, frequency of feeding solid/semi-solid foods[1], consumption of nutrient-dense food

[1]Data are not yet available, so variable construction and analysis for these data are not described in this version of the analysis plan.

groups yesterday and last week, food group diversity. We also assessed consumption of other fortified products (other than the project LNS) but preliminary analysis of the full sample showed that consumption of such products was extremely rare among iLiNS-DOSE infants, so no further analysis is planned.

Specific objectives of analysis

1.1 Primary objective

To compare infant and young child feeding practices and summary diet quality variables across intervention groups. Comparisons at baseline are descriptive, to assess comparability of groups. At all later time points, comparisons are to assess the effect of the intervention on IYCF practices.

1.2 Secondary objectives

To create summary diet quality variable(s) for potential use in analyses of main outcomes (effect modifiers)

To provide descriptive data on IYCF practices to contextualize results of the trials, and to aid readers in comparing to IYCF in other settings

To provide supporting descriptive data for full samples for manuscripts describing dietary intake of sub-samples, and for triangulating between food frequency and dietary data.

1.3 Exploratory analyses

Exploratory analyses will be described later, in separate SAP or in addenda to this SAP. Note that in the sister trial in Burkina Faso ("iLiNS-ZINC"), there is a pre-planned exploratory analysis: description of child feeding practices and factors associated with these practices including food security and seasonality, and relation with nutritional status (growth, iron status, morbidity).

3. Hypotheses to be tested

Stated qualitatively: Provision of LNS would not impact infant and young child feeding practices. More specifically, provision of LNS would not cause a change in:

- Breastfeeding (prevalence of any breastfeeding several time points throughout intervention, and reported frequency of breastfeeding the previous day)
- Frequency of feeding other solids foods (meals and snacks, or feeding episodes)[2];

[2]These data are not yet available and variable construction and analyses are not described in this version.

- Dietary diversity measured as food group diversity at or above the WHO cut-off[3];
- Number of nutrient-dense food groups (animal-source foods, fruits and vegetables)

4. Description of infant and young child feeding practice outcome variables

On the following pages, we present information on construction of outcome variables.

The table shows details on: data sources; variable names and variable construction; treatment of data and criteria for imputing missing values.

With few exceptions (detailed in the table) no data will be imputed for analyses of IYCF practices as outcomes.

However, in later analyses where longitudinal IYCF practice variables (e.g. indices summed across time points) may be used as potential effect modifiers, data could be imputed to avoid losing substantial numbers (those for whom data from all time points are not available).

Timing of outcomes: in the iLiNS-DOSE trial, target dates for collection of IYCF practices data were: Baseline (week 0, at ~5.5-6.5 mo of age); week 16; week 28; week 40; and endline (week 52). Data on breastfeeding practices, food group intake, and use of fortified products and vitamin/mineral drops were assessed at all time points. Data on frequency of meals and snacks were collected on KAP questionnaires at 6, 12, and 18 months of age (data not yet available).

[3] WHO (2008) Indicators for assessing infant and young child feeding practices: conclusions of a consensus meeting held 6–8 November 2007 in Washington D.C., USA.

Variable construction and handling for IYCF practice outcomes

Outcome	Variable name(s)	Location, variable construction	Criteria for errors, outliers	Criteria for transformation	Use of imputation
~9.6 mo: Still breastfed (%)	Source variables: • visitid • FFqStillFed Constructed variable: • stillbf	"visitid" is a numeric variable constructed from the string variable "NumberVisit" visitid=16 Still breastfed variable is on Form 13a & 13b, page 1; new constructed variable is named stillbf to harmonize with other sites	As above	No - dichotomous	Missing data can be coded as "1" (yes) if child is reported to be still breastfed at later time points.
~9.6 mo: Breastfed 6+ times yesterday (%)	Source variables: • Visitid • FFqNumBrtFeeds Constructed variable: • bf_6	"visitid" as above Frequency of breastfeeding from Form 13a & 13b, page 1 Coded "4" or "6" (different questionnaire versions had different codes, in error; syntax recodes all to same)	As above	No - categorical	No imputation for analysis of outcomes (group comparisons) .
~12 mo: Still breastfed (%)	As above	visitid=28 Else as above	As above	No - dichotomous	As above, can impute missing data if "yes" at later time points.
~12 mo: Breastfed 6+ times yesterday (%)	As above	visitid=28 Else as above	As above	No - categorical	No imputation

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~15 mo: Still breastfed (%)	As above	visitid=40 Else as above	As above	No - dichotomous	As above, can impute missing data if "yes" at later time points.
~15 mo: Breastfed 6+ times yesterday (%)	As above	visitid=40 Else as above	As above	No - categorical	No imputation
~18 mo: Still breastfed (%)	As above	visitid=52 Else as above	As above	No - dichotomous	No imputation; no later time points.
~18 mo: Breastfed 6+ times yesterday (%)	As above	visitid=52 Else as above	As above	No - categorical	No imputation

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~9.6 mo: 4+ food groups yesterday (WHO) (%)	• visitid • 1. Starchy staple groups: FqLiqPorr24, FFqThobwa24, FFqPorridge24, FFqOthGrain24, FFqTuber24 • 2. FFqLegume24 • 3. Dairy groups: FFqMilk24, FFqFormula24, FFqMilkTea24, FFqYogurt24, FFqDairy24 • 4. Flesh food groups: FFqOrgan24, FFqFlesh24, FFqFish24 FFqInsect24 • 5. FFqEggs24 • 6. Vit A-rich fr/vegs: FFqVitAVeg24, FFqLeaves24, FFqVitAFr24 • 7. Other fruits/veg: FFqOthFr24, FFqOthVeg24	Form 13, pages 2-3. For each of the 7 WHO food groups, a dichotomous variable is constructed, and coded "1" if any of the constituent groups=1; "0" if all are "0", missing if all are missing or combination of missing and 0. Constructed variables are: starch24, legume24, dairy24, flesh24, eggs24, vitafood24, othfrveg24 These new variables are summed to construct "dd24", a quais-continuous variable ranging from 0 to 7. A dichotomous variable is constructed for the WHO minimum diversity indicator, coded "0" if dd24<4, "1" if dd24 is 4, 5, 6 or 7, and missing if missing: dd24GE4	As above	No, dichotomous	If missing, variables for organ meats and insects are imputed to 0 as consumption was very rare (1.9% and 0.2% of child-days, respectively) No other imputation for analysis of outcomes (group comparisons).

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~9.6 mo: No ASF yesterday (%)	• visitid • FFqMilk24 • FFqFormula24 • FFqMilkTea24 • FFqYogurt24 • FFqOrgan24 • FFqFlesh24 • FFqEggs24 • FFqFish24 • FFqDairy24 • FFqInsects24	Form 13a, pages 2-3 The four dairy-containing fluids are coded, along with FFqDairy, into a new summary variable for any dairy yesterday "dairy24" A dichotomous ASF variable, for any ASF yesterday: "asf24yn". Coded "1" if dairy24 or any of organ, flesh, eggs or fish=1. Coded "0" if all are 0. Coded missing if all missing or combination of missing & 0. For this outcome, reporting % w/asf24=0.	As above; for solid foods, enumerator errors included not ticking "yes" when foods were circled, and not ticking "no" when not (i.e., not completing right-hand column on page 3). These values were corrected when missing or when errors observed, but we did not do a review of 100% of forms.	No - dichotomous	If missing, variables for organ meats and insects are imputed to 0 as consumption was very rare (1.9% and 0.2% of child-days, respectively) For all other ASF variables prevalence is >5% and there is no imputation.
~9.6 mo: Number of ASF food groups yesterday	• visitid • dairy24 • FFqOrgan24 • FFqFlesh24 • FFqEggs24 • FFqFish24 • FFqInsects24	Source of variables described above. The first 5 ASF variables are summed into a score, "asf24sum". A point is added for insects only if flesh foods are otherwise scored "0". Missing if any variables are missing.	As above	Quasi-continuous and non-normally distributed in full sample. No transformation planned.	As above.

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~9.6 mo: No fruits/vegetables yesterday (%)	• visitid • FFqVitAVeg24 • FFqLeaves24 • FFqVitAFr24 • FFqOthFr24 • FFqOthVeg24	Form 13a, page 3. A dichotomous fruit /vegetable variable, for any fr/veg yesterday: "frveg24yn". Coded "1" if any of the five groups=1. Coded "0" if all are 0. Coded missing if all missing or combination of missing and 0. For this outcome, reporting % with frveg24yn=0.	As above	No - dichotomous	No imputation for analysis of outcomes (group comparisons). In full sample, prevalence of child-days coded "1" is >10% for each of the 5.
~9.6 mo: Number of fruit/veg groups yesterday	As above	Source of variables described above. The 5 fruit/vegetable variables are summed into a score, "frveg24sum". Missing if any of the 5 are missing.	As above	Quasi-continuous and non-normally distributed in full sample. No transformation planned.	As above
~12 mo: 4+ food groups yesterday (WHO) (%)	As for ~9.6 month time point				
~12 mo: No ASF yesterday (%)					
~12 mo: Number of ASF food groups yesterday					
~12 mo: No fruits/vegetables yesterday (%)					
~12 mo: Number of fruit/veg groups yesterday					
~15 mo: 4+ food groups yesterday (WHO) (%)	As for ~9.6 month time point				
~15 mo: No ASF yesterday (%)					
~15 mo: Number of ASF food groups yesterday					

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~15 mo: No fruits/vegetables yesterday (%)					
~15 mo: Number of fruit/veg groups yesterday					
~18 mo: 4+ food groups yesterday (WHO) (%)	As for ~9.6 month time point				
~18 mo: No ASF yesterday (%)					
~18 mo: Number of ASF food groups yesterday					
~18 mo: No fruits/vegetables yesterday (%)					
~18 mo: Number of fruit/veg groups yesterday					
~9.6 mo: ASF score (out of 28)	• visitid • FFqMeat7 • FFqEggs7 • FFqFish7 • FFqDairy7	Form 13a, page 4. The 4 ASF variables are summed into a score, “asf7sum”. Range is 0 to 28. Missing if any are missing, or coded “66”, which means the child had the food, but respondent could not say how many days in last 7 days.	Obvious enumerator and data entry errors corrected; no other changes. Uncorrectable illegal codes coded missing.	Quasi-continuous and non-normally distributed in full sample. No transformation planned.	No imputation for analysis of outcomes (group comparisons).

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~9.6 mo: Lowest ASF tertile (%)	As above	Categorical variable "asf7T16" is constructed by creating tertiles of "asf7sum" based on the entire sample (all groups) at visit 16. The variable is coded 1-3 for lowest to highest tertile. Dichotomous variables are also generated, "i16_asfT1 – i16_asfT3" and coded 0 if no and 1 if yes for the indicated tertile. E.g. if a score is in the lowest tertile, i16_asfT1=1 and if not, i16_asfT1=0.	As above	No, constructed variables are categorical or dichotomous.	As above
~9.6 mo: Fruit/veg score (out of 35)	• Visitid • FFqVitAVeg7 • FFqLeaves7 • FFqVitAFr7 • FFqOthFr7 • FFqOthVeg7	Form 13a, page 4 The 5 fruit/vegetable variables are summed into a score, "frveg7sum". Range is 0 to 35. Missing if any are missing, or coded "66", which means the child had the food, but respondent could not say how many days in last 7 days.	Obvious enumerator and data entry errors corrected; no other changes. Uncorrectable illegal codes coded missing.	Quasi-continuous and non-normally distributed in full sample. No transformation planned.	No imputation for analysis of outcomes (group comparisons).

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~9.6 mo: Lowest fruit/veg tertile (%)	As above	Categorical variable "frveg7T16" is constructed by creating tertiles of "frveg7sum" based on the entire sample (all groups) at visit 16. The variable is coded 1-3 for lowest to highest tertile. Dichotomous variables are also generated, "i16_frvegT1 – i16_frvegT3" and coded 0 if no and 1 if yes for the indicated tertile. E.g. if a score is in the lowest tertile, i16_frvegT1=1 and if not, i16_frvegT1=0.	As above	No, constructed variables are categorical or dichotomous.	As above
~12 mo: ASF score (out of 28)	As for ~9.6 month time point				
~12 mo: Lowest ASF tertile (%)	As above	Constructed as above. New variables are asf7T28, and i28_asfT1 through i28_asfT3.	As above		
~12 mo: Fruit/veg score (out of 35)	As for ~9.6 month time point				
~12 mo: Lowest fruit/veg tertile (%)	As above	Constructed as above. New variables are frveg7T28, and i28_frvegT1 through i28_frvegT3.	As above		
~15 mo: ASF score (out of 28)	As for ~9.6 month time point				

Outcome	**Variable name(s)**	**Location, variable construction**	**Criteria for errors, outliers**	**Criteria for transformation**	**Use of imputation**
~15 mo: Lowest ASF tertile (%)	As above	Constructed as above. New variables are asf7T40, and i40_asfT1 through i40_asfT3.	As above		
~15 mo: Fruit/veg score (out of 35)	As for ~9.6 month time point				
~15 mo: Lowest fruit/veg tertile (%)	As above	Constructed as above. New variables are frveg7T40, and i40_frvegT1 through i40_frvegT3.	As above		
~18 mo: ASF score (out of 28)	As for ~9.6 month time point				
~18 mo: Lowest ASF tertile (%)	As above	Constructed as above. New variables are asf7T52, and i52_asfT1 through i52_asfT3.	As above		
~18 mo: Fruit/veg score (out of 35)	As for ~9.6 month time point				
~18 mo: Lowest fruit/veg tertile (%)	As above	Constructed as above. New variables are frveg7T52, and i52_frvegT1 through i52_frvegT3.	As above		

5. Approach to analysis and exclusions specific to this analysis

All tests will be two-sided, at 5% level of significance.

Since varying numbers of observations are available depending on the time point (i.e., there were a substantial number of missed visits), sample sizes by group will be reported for each time point. If specific outcome variables are missing for more than 10% of infants (with denominator being total records available for the time point) we will report the number of observations used per specific outcome analysis.

First analysis will be by intention-to-treat, but excluding observations that are more than ±4 weeks (28 days) from the median observed age for visits 0, 16, 28 and 40. That is, at time points, 0, 16, 28 and 40, child age could vary by up to 8 weeks.

Exclusions for visit 52 (endline) will be handled differently. For visit 52, we will exclude visits that are more than 28 days before the target date as for other time points. But since many endline visits were late, and since feeding practices change more slowly by 18 months of age compared to earlier time points, we will include visits up to 6 weeks after the target date for the child. Rationale: visit 52 was planned to occur exactly one year after the date of enrollment. However, so long as visit 52 occurred within one month of the target date, the child received a 2 week supply of LNS. So, if the endline FFQ occurred within ~6 weeks (42 days) of target, the child should have had LNS in the week prior to the FFQ.

Data on subjects who were lost to follow-up (either temporarily or permanently) will be included in the analysis for all time points where data are available.

After the intent-to-treat analysis, we will perform a per protocol analysis as follows: on the FFQ (Form 13) questionnaire itself, respondents were asked how many days in the last 7 days the child consumed LNS the household received from the project. In preliminary analysis, it is apparent that there is heavy data lumping on “0” and “7”. Per protocol analysis will compare feeding practices for infants who were reported to consume LNS 7 days in the last 7 days, as compared to those who were either in the delayed intervention group, or reported to consume the supplement 0 days. Infants who were reported to consume LNS 1-6 days in the last 7 days will be excluded from this analysis (0.3-7.9% of infants, depending on time point).

After group codes are revealed, if at any time point fewer than 10% of those in the LNS groups were reported to consume on 0 days in the last 7 days, this analysis will not be performed.

6. Statistical methods

5.1 Software

All analyses will be done using SAS version 9.3 (SAS Inst. Cary, NC, USA) or Stata version 10.1 or higher (StataCorp, TX, USA). The WHO 2006 Child Growth Standards will be used for age-and-sex standardization of child weight, length, head circumference, arm circumference and weight-for-height.

5.2 Sample size and attrition

Sample sizes by group will be presented for each time point (**Table 1 and Figure 1**), and differential attrition will be assessed with chi-square tests at each time point.

5.3 Background characteristics

Selected background characteristics (measured at baseline) will be examined by group for baseline and endline samples (**Table 2**). Feeding practices at baseline will be presented, by group (**Table 3**).

5.4 Analysis of the effect of the intervention

General comments:

a. The hypothesis stated in section 3 is a non-equivalence hypothesis. However, the study was not powered for IYCF practices outcomes and we are severely underpowered for equivalence analyses, particularly for dichotomous outcomes (such as prevalence of continued breastfeeding at any time point). Therefore the more traditional approach in the nutrition literature of analyzing for significant differences will be followed in the first instance. This limitation will be clearly explained in the discussion section of any publication.
b. For quasi-continuous variables, we will supplement this with an equivalence approach to hypothesis testing, to help inform conclusions from this analysis.

Analysis of the effect of the intervention will follow these steps:

a. A set of pre-specified potential covariates will be examined through reviewing correlations and collinearity (e.g. using "collin" command in Stata). Variables with VIF > 10 will be assessed and a reduced set of variables will be retained, such that all VIF are < 10.
b. We will test the null hypothesis of no difference among the four treatment groups (iLiNS-DOSE) using ANCOVA or logistic regression, and controlling for the pre-specified covariates (most of which are also potential effect modifiers, see below).
c. For all analyses, if the global null hypothesis is rejected at 0.05 level, then we will perform post-hoc pairwise comparisons of all four groups using appropriate

adjustments for multiple comparisons and Scheffe's test to examine other contrasts of interest.

d. The effects of potential effect modifiers will be assessed with an interaction term in the ANCOVA or logistic regression model. Each interaction will be assessed separately.
e. Significant interactions ($p < 0.05$) will be further examined with stratified analyses, estimation of separate regression lines, or estimation of adjusted means at key points of the covariate, in order to understand the nature of the effect modification.
f. For quasi-continuous outcomes (number of fruit/vegetable groups consumed yesterday; number of animal-source food groups consumed yesterday, and fruit/vegetable and animal-source food scores for last week) equivalence will be assessed based on defined margins. Margins for yesterday will be ±1.0 (one more or one fewer fruit vegetable group; one more or one fewer animal-source food group). For scores for last week, the margin will be ±5 points for the fruit/vegetable scores and ±4 for animal-source foods (~= a difference of one group/day in each).
g. We will assess equivalence in the context of ANCOVA models, controlling for the same pre-specified covariates as noted above. Equivalence will be determined to exist if the 90% confidence interval for the difference between the means is entirely contained within the negative and the positive values of the equivalence margin.
h. Confidence intervals will be adjusted for multiple comparisons.
i. For each outcome, if results are inconclusive, all LNS groups will be combined and 2-group comparisons will be made, to improve power to detect differences.

5.5 Covariates in main effects models

In theory, a variety of community-, household-, maternal-, and child-level characteristics could affect child feeding practices independently of the intervention. Data are available for the covariates listed below. All covariates are as measured at baseline, with the exception of child age. Since child age at each visit can vary, child age at time of measure will be included in models. Before making final decisions on inclusion of covariates, completeness of data for the covariates will be considered and covariates will be excluded if loss of sample size is judged too large.

- Month of study during which the child was enrolled
- Characteristics of households
    - Location of household, or distance to market and distance to clinic

  - Baseline HH food security (HFIA score)
  - Baseline HH asset score
  - Baseline HH housing quality score
  - Baseline HH livestock assets score
  - Number of underfives in the HH, assessed at baseline
- Maternal education
- Child's characteristics
  - Child age
  - Child sex
  - Child length-for-age z-score (LAZ) at baseline
  - Child weight-for-length z-score (WLZ) at baseline

5.6 List of potential effect modifiers to be examined

Similarly most of the covariates identified above could interact with the provision of LNS to produce differential effects on feeding practices. Specifically, the household, maternal, and child-level characteristics, but not the temporal variable, will be evaluated for their potential to interact with intervention group.

5.7 Future exploratory (path) analyses

There are a number of potential effect modifiers not yet available, but which could interact with provision of LNS to produce differential effects. These may be included in later exploratory analyses:

- Household
  - Proxy for income (expenditures)
  - Grandmother lives on compound/with the child
- Maternal
  - Index for responsive feeding (composite from KAP)
  - Maternal decision-making power (composite from KAP)
  - Maternal depression
- Child
  - Morbidity

- Appetite

7. Design of tables and figures

See following pages for a list of tables and example figures that will be examined by the manuscript writing group:

Table 1. Number of observations and missing visits, by intervention group and by time point

Table 2. Background characteristics of study participants, baseline and endline samples, by intervention group

Table 3. Feeding practices at enrollment, by intervention group

Table 4a. Continued breastfeeding and frequency of breastfeeding yesterday, by intervention group

(Table 4b. Continued breastfeeding and frequency of breastfeeding yesterday, pairwise comparisons. This table, with pairwise comparisons, will be prepared if overall tests are significant)

Table 5a Infant diet quality yesterday: food group diversity & nutrient-dense food groups by intervention group

(Table 5b Infant diet quality yesterday: food group diversity & nutrient-dense food groups, pairwise comparisons. This table, with pairwise comparisons, will be prepared if overall tests are significant)

Table 6a Infant diet quality last week: nutrient-dense food groups, by intervention group

(Table 6b Infant diet quality last week: nutrient-dense food groups, pairwise comparisons This table, with pairwise comparisons, will be prepared if overall tests are significant)

Example figures (final set of figures to be determined):

Figure 1 Participant flow

Figure 2 Breastfeeding practices, by intervention group

Figure 3 Nutrient-dense food groups yesterday, by intervention group

Figure 4 Possibly, a Figure illustrating comparison between adjusted confidence intervals and equivalence margins

Table 1. Number of observations and missing visits, by time point (FFQ)

Age[a]	X.X mo				X.X mo				XX.X mo				XX.X mo				XX.X mo			
Group[b]	0	10	20	40	0	10	20	40	0	10	20	40	0	10	20	40	0	10	20	40
Number permanently lost to follow-up																				
Number missing																				
% missing																				
Number completed																				
Total[c]																				

[a] Median observed age. Infants were enrolled at 5.5-6.5 months of age. Subsequent data collection time points for infant feeding practices were: 16 weeks, 28 weeks, 40 weeks, and 52 weeks of participation in the study.

[b] Group defined by assignment to receive 0 g LNS (control; delayed intervention); 10 g; 20 g; or 40 g.

[c] Total number of children who should have been in dataset at this time point (excluding permanent loss to follow-up).

Table 2. Background characteristics of study participants

	N	(missing)	Control	LNS[a] 10 g	LNS 20 g	LNS 40 g	All	P-value[b]
Time of enrollment								
Baseline sample								
Endline sample								
Location of HH (or distances to clinic, market)								
Baseline sample								
Endline sample								
HFIA score at baseline								
Baseline sample								
Endline sample								
HH asset score at baseline								
Baseline sample								
Endline sample								
HH housing quality score at baseline								
Baseline sample								
Endline sample								
HH livestock assets score at baseline								
Baseline sample								
Endline sample								
Number of underfives in the HH at baseline								
Baseline sample								
Endline sample								
Maternal education (in y)								
Baseline sample								
Endline sample								
Child age at baseline (in mo)								
Baseline sample								
Endline sample								
Child sex (% female)								
Baseline sample								
Endline sample								
Child length-for-age z-score (LAZ) at baseline								
Baseline sample								
Endline sample								

Child weight-for-length z-score (WLZ) at baseline

Baseline sample

Endline sample

[a] LNS=lipid-based nutrient supplement; the quantities indicated are the size of the infant's daily dose, delivered to the household *[indicate weekly or bi-weekly]*.

[b] Comparison between intervention groups at each time point; p-value for (describe tests).

Table 3. Feeding practices at enrollment, by intervention group

	N[a]	(missing)	Control	LNS[b] 10 g	LNS 20 g	LNS 40 g	All	P-value[c]
Exclusive breastfeeding, infants <6 mo[d]								
Frequency of breastfeeding[e]								
None								
6+ times								
Food groups consumed by the infant yesterday[f]								
Porridge								
Other grain-based foods or roots/tubers								
Any legumes or nuts								
Any dairy (liquids, semi-solids, solids)								
Any meat, poultry or fish								
Any eggs								
Any vitamin A-rich fruit and/or vegetable								
Any other fruit and/or vegetable								

[a] Number of infants at enrollment. For exclusive breastfeeding n is the number of infants who were less than 6 months of age at enrollment.

[b] LNS=lipid-based nutrient supplement; the quantities indicated are the size of the infant's daily dose, delivered to the household *[indicate weekly or bi-weekly]*.

[c] Comparison between intervention groups; p-value for ANOVA (continuous and quasi-continuous variables) or chi-square test (categorical variables).

[d] Exclusive breastfeeding defined based on negative responses to a series of questions on fluids, semi-solids and solids consumed yesterday.

[e] Respondents were asked if the baby was breastfed yesterday and if so they were read the following options: Only at night; only 1 or 2 times during the day; about 3 to 5 times during the day; at least 6 times during the day.

[f] Respondents were asked (yes/no) if the infant was given any of a list of fluids and semi-solids yesterday, and solid food consumption was determined using an open, guided recall of the previous day (qualitative 24-hour recall). Foods reported in the recall were circled and coded into groups.

Table 4a. Frequency of breastfeeding yesterday, by intervention group

	N[a]	(missing)	Control	LNS[b] 10 g	LNS 20 g	LNS 40 g	All	P-value[c]
Frequency of breastfeeding[d]								
At ~ X.X mo[e] (%)								
None								
6+ times								
At ~ XX.X mo (%)								
None								
6+ times								
At ~ XX.X mo (%)								
None								
6+ times								
At ~ XX.X mo (%)								
None								
6+ times								

[a] Number of infants not permanently lost to follow-up at each age/time point.

[b] LNS=lipid-based nutrient supplement; the quantities indicated are the size of the infant's daily dose, delivered to the household *[indicate weekly or bi-weekly]*.

[c] Values presented are unadjusted prevalences. Statistical tests are for adjusted analyses: logistic regression, controlling for……

[d] Respondents were asked if the baby was breastfed yesterday and if so they were read the following options: Only at night; only 1 or 2 times during the day; about 3 to 5 times during the day; at least 6 times during the day.

[e] Approximate age shown is median observed age. After enrollment, additional data collection time points for infant feeding practices were: 16 weeks, 28 weeks, 40 weeks, and 52 weeks of participation in the study.

Table 5a Infant diet quality yesterday: food group diversity & nutrient-dense food groups

	N[a]	(missing)	Control	LNS[b] 10 g	LNS 20 g	LNS 40 g	All	P-value[c]
Food groups consumed by the infant yesterday[d]								
At ~ X.X mo[e] (%)								
4+ food groups (WHO indicator[f], %)								
No ASF[g] (%)								
Mean or median # ASF groups (of 5)[h]								
No fruits/vegetables (%)								
Mean or median # fruit/vegetable groups (of 5)[i]								
At ~ XX.X mo (%)								
4+ food groups (WHO indicator, %)								
No ASF (%)								
Mean or median # ASF groups (of 5)								
No fruits/vegetables (%)								
Mean or median # fruit/vegetable groups (of 5)								
At ~ XX.X mo (%)								
4+ food groups (WHO indicator, %)								
No ASF (%)								
Mean or median # ASF groups (of 5)								
No fruits/vegetables (%)								
Mean or median # fruit/vegetable groups (of 5)								
At ~ XX.X mo (%)								
4+ food groups (WHO indicator, %)								
No ASF (%)								
Mean or median # ASF groups (of 5)								
No fruits/vegetables (%)								
Mean or median # fruit/vegetable groups (of 5)								

[a] Number of infants not permanently lost to follow-up at each age/time point.

[b] LNS=lipid-based nutrient supplement; the quantities indicated are the size of the infant's daily dose, delivered to the household *[indicate weekly or bi-weekly]*.

[c] Values presented are unadjusted means (SD) or medians (I-Q ranges) and prevalences. Decision on presenting means or medians will be made after examination of distributions. Statistical tests are for adjusted analyses; analysis of covariance and logistic regression, controlling for…

[d] Respondents were asked (yes/no) if the infant was given any of a list of fluids and semi-solids yesterday, and solid food consumption was determined using an open, guided recall of the previous day (qualitative 24-hour recall). Foods reported in the recall were circled and coded into groups.

[e] Approximate age shown is median observed age. After enrollment, additional data collection time points for infant feeding practices were: 16 weeks, 28 weeks, 40 weeks, and 52 weeks of participation in the study.

[f] The WHO indicator sums seven food groups and a score of 4 or more of 7 is associated with higher nutrient density (WHO, 200X); the food groups are: 1) grains, roots and tubers; 2) legumes and nuts; 3) dairy products; 4) flesh foods; 5) eggs; 6) vitamin-A rich fruits and vegetables; and 7) other fruits and vegetables.

[g] ASF=animal-source food.

[h] The 5 ASF groups are: 1) organ meats; 2) other meat/poultry; 3) fish; 4) eggs; and 5) dairy.

[i] The 5 fruit and vegetables groups are: 1) vitamin A-rich orange/yellow vegetables; 2) dark green leafy vegetables; 3) other vegetables; 4) vitamin A-rich fruits; and 5) other fruits..

Table 5a Infant diet quality last week: nutrient-dense food groups, by intervention group

	N[a]	(missing)	Control	LNS[b] 10 g	LNS 20 g	LNS 40 g	All	P-value[c]
Food groups consumed by the infant in the last 7 days[d]								
At ~ X.X mo[e] (%)								
Mean or median ASF score (of 28)[g]								
Lowest ASF tertile (%)[h]								
Mean or median fruit/vegetable score (of 35)[i]								
Lowest fruit/vegetable tertile (%)								
At ~ XX.X mo (%)								
Mean or median ASF score (of 28)								
Lowest ASF tertile (%)								
Mean or median fruit/vegetable score (of 35)								
Lowest fruit/vegetable tertile (%)								
At ~ XX.X mo (%)								
Mean or median ASF score (of 28)								
Lowest ASF tertile (%)								
Mean or median fruit/vegetable score (of 35)								
Lowest fruit/vegetable tertile (%)								
At ~ XX.X mo (%)								
Mean or median ASF score (of 28)								
Lowest ASF tertile (%)								
Mean or median fruit/vegetable score (of 35)								
Lowest fruit/vegetable tertile (%)								

[a] Number of infants not permanently lost to follow-up at each age/time point.

[b] LNS=lipid-based nutrient supplement; the quantities indicated are the size of the infant's daily dose, delivered to the household *[indicate weekly or bi-weekly]*.

[c] Values presented are unadjusted means (SD) or medians (I-Q ranges) and prevalences. Decision on presenting means or medians will be made after examination of distributions. Statistical tests are for adjusted analyses; analysis of covariance and logistic regression, controlling for ….

[d] Respondents were asked how many days in the last seven days the infant was given any of a list of foods. Foods were read to the respondent in groups. The respondent was also instructed to consider ingredients of mixed dishes.

[e] Approximate age shown is median observed age. After enrollment, additional data collection time points for infant feeding practices were: 16 weeks, 28 weeks, 40 weeks, and 52 weeks of participation in the study.

[f] ASF=animal-source food.

[g] The ASF score equals the sum, across four animal-source food groups, of the number of days in the last seven days the respondent reported the infant consumed a food in the group. The four groups are: 1) meat/poultry; 2) fish; 3) eggs; and 4) dairy. Scores could range from 0 to 28.

[h] Tertiles are based on the full sample at each time point. The table shows the percent of children in each intervention group falling into the lowest tertile for the ASF score and for the fruit/vegetable score.

[i] The fruit/vegetable score equals the sum, across five fruit and vegetable food groups, of the number of days in the last seven days the respondent reported that the child consumed a food in the group. The five groups are: 1) vitamin A-rich yellow/orange vegetables; 2) dark green leafy vegetables; 3) vitamin A-rich fruits; 4) other fruits; 5) other vegetables. Scores could range from 0 to 35.

Figure 1

Figure 1 will present a detailed participant flow chart (CONSORT diagram). The Figure will include numbers and reasons for permanent and temporary loss-to-follow-up at each time point where outcomes are evaluated (visits at weeks 16, 28, 40, and 52).

Figure 2 Breastfeeding practices, by intervention group

Either bar or line graphs, with confidence intervals, will be used to present adjusted or unadjusted prevalences by intervention group and by time point. If relevant, the Figure legend will indicate covariates adjusted for. Data presented here are illustrative only (observations at each time point were randomly allocated to 4 groups).

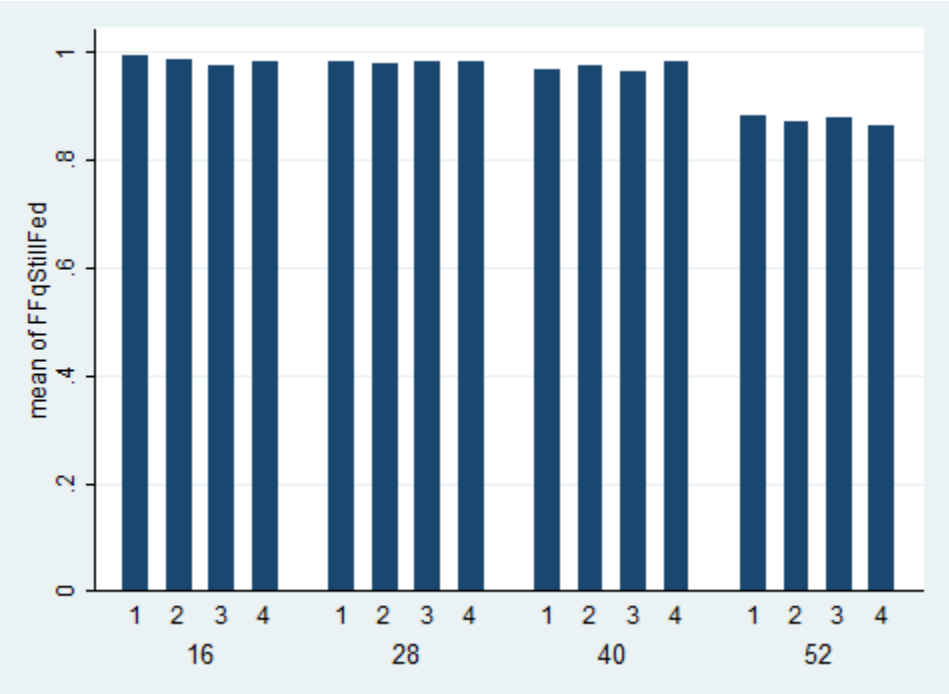
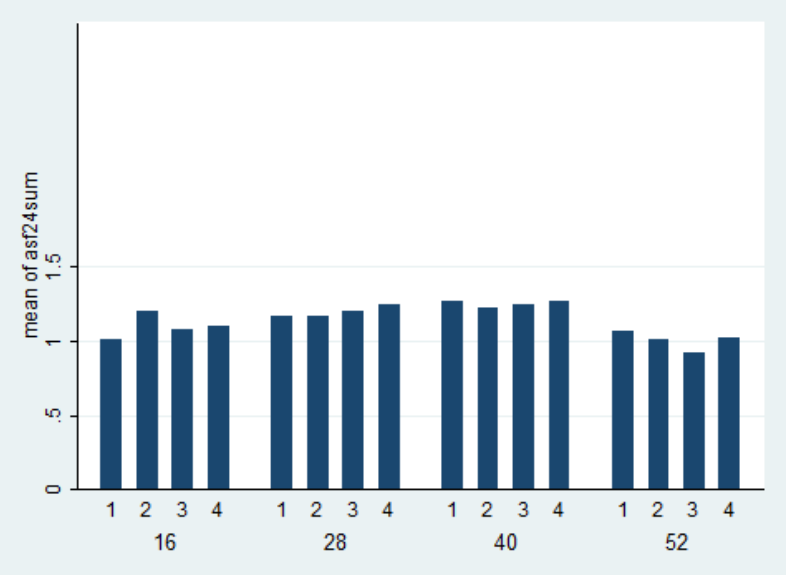


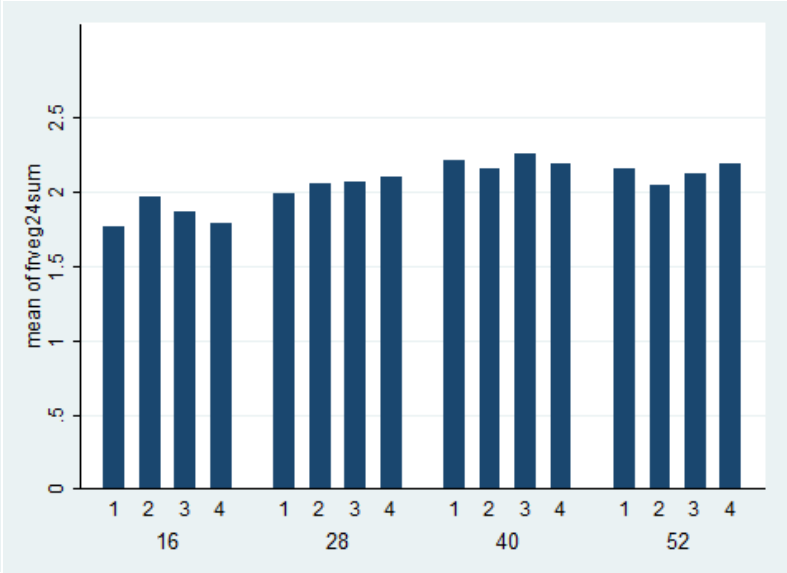
Figure 3 Mean number of nutrient-dense food groups yesterday, by intervention group

Either bar or line graphs, with standard error bars, will be used to present adjusted or unadjusted means by intervention group and by time point. If relevant, the Figure legend will indicate covariates adjusted for. Data presented here are illustrative only (observations at each time point were randomly allocated to 4 groups).

Panel A – Animal-source food groups



Panel B – Fruit/vegetable food groups



Prevention of Linear Growth Faltering in Infants and Young Children With Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical Analysis Plan

Appendix 06: Hypothetical Willingness-to-Pay for LNS and Likuni Phala at Baseline (added on 23.04.2014)

Contents

1. Brief Introduction and Motivation

In this paper we will present baseline estimates of hypothetical willingness-to-pay (WTP) for a small-quantity preventative lipid-based nutrient supplement (LNS) product formulated for consumption during early childhood, from approximately 6-24 months. Using contingent valuation methods, we elicited hypothetical WTP for a week's supply of LNS from households participating in the iLiNS-DOSE randomized control nutrition trial in Malawi. As a comparator, we also elicited hypothetical WTP for a week's supply of Likuni Phala (LP), a familiar, locally-available product commonly in this iLiNS study area. For both LNS and LP, after eliciting WTP for a week's supply of the product, we used a set of follow-up questions to assess hypothetical WTP in the long-term (i.e., WTP for a week's supply regularly over the coming year).

Preventative LNS products are intended to be consumed daily for many months as a supplement to breast milk and traditional foods (Dewey and Arimond 2012; Nutriset 2011). This is in contrast to ready-to-use therapeutic foods such as Plumpy'Nut©, which are primarily used in emergency settings and are administered in relatively large doses over a short period of time to treat children with severe acute malnutrition. While the international donor community has historically purchased and distributed therapeutic nutritional products for severely malnourished children for free via public channels, the differences in usage of preventative LNS products coupled with the potentially large and heterogeneous population of women and children who may benefit from them will make full subsidization of preventative LNS products much more expensive and less likely (Lybbert 2012). Thus, a hybrid distribution system that reaches target consumers through both public channels and retail markets may be recommended.

In this hybrid setting, in addition to the opportunity costs associated with procuring and consuming preventative LNS products, some households may also be required to pay for them. Our estimates of willingness-to-pay (WTP) for LNS will shed light on household valuation of LNS and the factors that influence WTP.[1] Moreover, our data on WTP for LP will provide a benchmark from which we can evaluate WTP for LNS relative to a familiar, locally-available product. This collection of results will provide a starting point for characterizing demand for LNS, which in turn may guide policy decisions regarding the price LNS consumers might be expected to pay as well as help establish a targeting mechanism to distribute LNS.

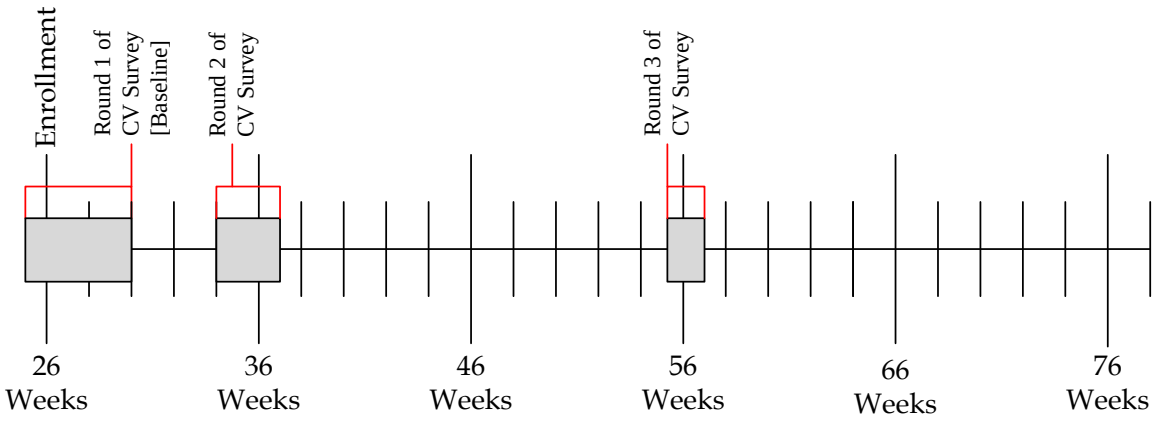
2. Description of Variables

The following sections describe the dependent and explanatory variables that will be used to model WTP. Note that the baseline contingent valuation survey was to be administered within a

[1] The randomized trial is evaluating the efficacy of LNS for childhood consumption.

few weeks of enrollment.[2] Figure 1 depicts the relationship between infant enrollment into the randomized trial and the actual timing of each round of the contingent valuation survey.[3] Time is measured in weeks from the birth of the child. The grey boxes indicate the approximate range of time when enrollment and contingent valuation surveys were administered.

FIGURE 1: TIMELINE OF iLiNS-DOSE INTERVENTION AND CONTINGENT VALUATION (CV) SURVEY



2.1 Dependent Variables

- WTP for a week's supply of LNS at baseline in 4th quarter 2011 US dollars.
- WTP for a week's supply of LP at baseline in 4th quarter 2011 US dollars.
- Difference in WTP for a week's supply of LNS and LP at baseline in 4th quarter 2011 US dollars.
- Long-term (i.e., one year) WTP for a week's supply of LNS at baseline in 4th quarter 2011 US dollars.
- Long-term (i.e., one year) WTP for a week's supply of LP at baseline in 4th quarter 2011 US dollars.
- Difference in long-term (i.e., one year) WTP for a week's supply of LNS and LP at baseline in 4th quarter 2011 US dollars.

Note: The distributions of WTP for LNS and LP are right-skewed. To account for this in our models, we may transform WTP to ln(WTP).[4]

[2] In some instances, contingent valuation surveys were administered a few weeks past the planned enumeration date due to logistical reasons and difficulty locating respondents.
[3] The focus of this manuscript will be baseline hWTP only.
[4] Because the natural log of zero is undefined, we will set all zero WTP values to a value slightly smaller than the minimum non-zero value of ln(WTP).

2.2 Explanatory Variables

Childhood consumption of LNS may have private benefits that accrue to the iLiNS child and her household at different points along the lifecycle. The immediate- and short-term benefits potentially include reduced child morbidity (Martorell 1999; Allen and Gillespie 2001), which may decrease household expenditures on health care and ease the household's time and, perhaps, budget constraints by freeing up maternal time spent caring for a sick child. In the long-term, the household may benefit from improvements in the child's physical capacity, cognitive ability, and accumulation of human capital, leading to productivity gains in adulthood (Alderman 2010), thus increasing the household's incentive to invest in early childhood health.

There may also be costs associated with childhood consumption of LNS, such as the time spent procuring and consuming LNS or any unpleasant side-effects associated with its consumption. Given households' preferences and constraints, a household's expected stream of benefits (which may be shaped by characteristics such as level of education, demographic composition of the household, discount rate, and child and maternal health) coupled with the costs associated with consuming LNS will influence the private value (WTP) for LNS. The expected relationship between WTP and the following respondent, household, maternal characteristics, and child characteristics will be tested using Ordinary Least Squares (OLS) models and described in Section 3 below.[5]

Respondent Baseline Characteristics:

- Head of Household: Indicator variable that = 1 if the respondent is the iLiNS head of household and = 0 if respondent is the primary caregiver for the iLiNS child.[6]
- Age: Respondent's age in years.
- Education: Number of completed years of formal education by the respondent.

Household Baseline Characteristics:

- Children Under Five: The number of children under five years of age who are household members at baseline.[7]

[5] In some cases, the relationship between WTP and a covariate may be non-linear. In particular, respondent age, respondent income, household food security, and household expenditures may have an inverted u-shaped relationship with WTP, where WTP is lower at the tails of the covariate distribution. To account for this potential non-linearity, we will also include squared terms.
Note that some of the variables included in this list (and any variant of them, including squared terms and interactions) may be too highly correlated to include both in the model. We will test all independent variables for correlation and omit those deemed to be too highly correlated.

[6] The respondent to the contingent valuation survey was determined randomly (by tossing a coin) to be either the head of the household or the iLiNS child's primary caregiver. In cases where the caregiver is also the head of household, this variable is coded as = 1 (head of household).

[7] Household members are defined as people who have been regularly sleeping in the same dwelling and sharing food from the same cooking pots for at least the last three months.

- Percent Under Five: The percentage of household members who are under five years of age at baseline, defined as (children under five/household size)*100.
- HFIA Score: The Household Food Insecurity Access (HFIA) Score is a continuous measure of the degree of food insecurity in the household. For each of nine questions, the survey respondent, who is the person primarily responsible for food preparation and meals in the household, indicates whether anyone in the household experienced the food insecurity condition in the previous four weeks. If yes, the respondent indicates how frequently the specific condition was experienced, where 'rarely' = 1-2 times in the past four weeks, 'sometimes' = 3-10 times in the past four weeks, and 'often' = more than 10 times in the past four weeks. Each household receives a score from 0-27 based on a simple sum of the frequency of occurrence of each food insecurity condition, where 'never' = 0, 'sometimes' = 2 points, and 'often' = 3 points. The higher the score, the higher the degree of household food insecurity experienced in the previous four weeks.
- Household Asset Index: A proxy measure of household socioeconomic status based on ownership of a set of assets (radio, television, refrigerator, cell phone, and stove), lighting source, drinking water supply in the dry season, sanitation facilities, and flooring materials. Household ownership of this set of assets is combined into an index (with a mean of zero and a standard deviation of one) using principal components analysis. Higher asset index scores indicate relatively 'better-off' households.
- Household Per Capita Expenditures: Total daily per capita (PC) expenditures, composed of non-food expenditures plus food expenditures (which includes the value of purchased and home-produced foods) in 4th quarter 2011 US dollars.
- Percent Food Expenditures: The percentage of total daily per capita expenditures that go toward food, defined as (PC daily food expenditures/PC total daily expenditures)*100.
- Discount Rate: Relative measure of the household's discount rate determined by playing a game at baseline in which a respondent was shown two equal-sized tins of rice and was then asked to measure out the quantity (from 0-10) of rice into a third tin that would make him/her indifferent between receiving the first tin of rice alone in a week and the second tin plus the additional amount measured into the third tin in one month.[8]
- Risk Behavior: The measure of relative household risk aversion was generated by playing a game at baseline in which a respondent was given 150 Malawian Kwacha (approximately 0.38 USD) and allowed to bet a portion of the mount flipping a coin. If the coin landed on heads, the respondent lost half of the amount bet. If the coin landed on tails, the respondent gained the amount bet.

[8] To determine whether the respondent received rice in a week or a month, s/he rolled a 10-sided die. If the number rolled was smaller than the amount of rice measured, the first tine of rice alone was delivered to the respondent in a week, and if the number rolled was equal to or greater than the amount of rice measured, the second tin of rice plus the amount measured into the third tin was delivered to the respondent in a month. The quantity of additional rice measured into the third tin by the respondent serves as his/her individual discount rate relative to the rest of the sample.

Maternal Baseline Characteristics

- Maternal Height: Mother's height in meters measured at enrollment.[9]
- Maternal BMI: Mother's body mass index at enrollment.
- Age: Mother's age in years.
- Education: Number of completed years of formal education by the iLiNS child's mother.

Child Baseline Characteristics

- LNS: Dummy variable = 1 if iLiNS child randomized to receive LNS and = 0 if iLiNS child randomized to receive delayed intervention.[10]
- Child's Height-for-Age Z-score: Child's height-for-age Z-score measured at enrollment.[11]
- Primiparity: Dummy variable = 1 if iLiNS child is mother's first child.
- Male: Dummy variable = 1 if iLiNS child is male and = 0 if iLiNS child is female.

Other Covariates/Controls

- Month: Dummy variables indicating the month the baseline contingent valuation survey was administered.
- Year: Dummy variables indicating the year the baseline contingent valuation survey was administered.
- Enumerator: Set of enumerator control variables.
- Language of Enumeration: Dummy variable = 1 if language of enumeration is Chewa and = 0 if language of enumeration is Yao.

3. Statistical Methods

3.1 Data Cleaning

Cleaning of the SES data follows the same procedure outlined in the main analysis plan with the research assistant generating queries and the SES Coordinator resolving the queries.

[9] The perceived importance of maternal height may be relative to the height of other women in the iLiNS study catchment area. As such, we may also normalize height by the average height of other women in the iLiNS-DOSE trial.

[10] We may also estimate the models using a set of LNS treatment group dummy variables to assess whether there is any statistically significant difference in WTP, all else equal, across the treatment arms.

[11] The perceived importance of child height-for-age may be relative to the height-for-age of other children in the iLiNS study catchment area. As such, we may also normalize child height-for-age by the average height-for-age of children in the iLiNS-DOSE trial.

3.2 Outliers

Identification and treatment of outliers in the SES data and maternal and child nutrition variables will follow the treatment described in the main statistical plan.

3.3 Software

All statistical analyses will be performed with Stata 13 statistical package.

3.4 Analysis

3.4.1 Summary Baseline Characteristics

Summary statistics, including mean (count for dichotomous variables), standard deviation (percentage for dichotomous variables), minimum, and maximum for all explanatory variables will be presented in Table 1. As a check for the success of randomization, we will report differences in mean explanatory variables across treatment groups.

3.4.2 Summary of Short- and Long-term WTP

Summary statistics, including mean, standard deviation, minimum, and maximum for short-term (i.e., a week's supply) WTP for LNS, LP, and the difference in short-term WTP between the two products will be presented in Table 2. Table 3 will presented short-term WTP across treatment groups and respondents.

Tables 4 and 5 will present the same summary statistics but for long-term (i.e., one year) WTP for LNS, LP, and the differences in long-term WTP between the two products.

3.4.3 Factors Associated with WTP

Regression results will be presented in Table 6 (short-term WTP) and Table 7 (long-term WTP). We will use ordinary least squares (OLS) to estimate the relationship between baseline WTP for LNS and a set of characteristics that, based on theory and previous empirical work, we expect to be associated with WTP.[12]

For i = 1, 2, …, N contingent valuation survey respondents and m = 1, 2, …, M iLiNS children, we will estimate $WTP_i = \alpha + x_i'\beta_x + H_i'\beta_h + N_m'\beta_n + C_i'\beta_c + u_i$,[13] where WTP_i is respondent i's stated maximum WTP for a week's supply of LNS, x_i, is a vector of respondent baseline

[12] If WTP is censored at zero—that is, WTP is actually negative (and unobserved) for some respondents who would require payment to take LNS/LP—OLS may lead to inconsistent estimates (Cameron and Trivedi 2005). A tobit model can be used to account for censoring but is not without tradeoffs. The tobit model relies on normally distributed and homoscedastic errors for consistency, and since we observe only a small proportion of zeroes in our data, we opt for OLS over a tobit specification.

[13] In cases where the iLiNS mother was the respondent to the valuation survey, the respondent denoted j, is also the iLiNS woman, denoted m.

socioeconomic characteristics, H_i is a vector of respondent i's household baseline characteristics, N_m, is a vector of maternal and child baseline characteristics including indicators of maternal and child nutritional status, C_i is a vector of other control variables, and u_i is the error term. We will estimate a parallel model for baseline WTP for LP.

We will also use OLS to estimate the factors associated with the difference in WTP for LSN and LP at baseline, defined as $WTP^{LNS}_i - WTP^{LP}_i$. This will be modeled as $WTP^{LNS}_i - WTP^{LP}_i = \alpha + x_i'\beta_x + H_i'\beta_h + N_m'\beta_n + C_i'\beta_c + u_i$.

4. References

Alderman, Harold. 2010. "The Economic Cost of a Poor Start to Life." *Journal of Developmental Origins of Health and Disease*. 1(1): 19-25.

Allen, Lindsey and Stuart Gillespie. 2001. "What Works? A Review of the Efficacy and Effectiveness of Nutrition Interventions." Nutrition and Development Series Five. Manila: Asian Development Bank.

Dewey, Kathryn G. and Mary Arimond. 2012. "Lipid-Based Nutrient Supplements: How Can They Combat Child Malnutrition?" PLoS Med. 9(9): e1001314.

Lartey, Anna. 2008. "Maternal and Child Nutrition in Sub-Saharan Africa: Challenges and Interventions." *Proceedings of the Nutrition Society*. 67:105-108.

Lybbert, Travis. 2012. "Hybrid Public-Private Delivery of Preventative Lipid-Based Nutrient Supplement Products: Key Challenges, Opportunities, and Players in an Emerging Product Space." *SCN News*. UN Standing Committee on Nutrition. 39: 32-39.

Martorell, Reynaldo. 1999. "The Nature of Child Malnutrition and its Long-Term Implications." *Food and Nutrition Bulletin*. 20(3): 288-292.

Nutriset. 2011. "Prevention of Malnutrition." http://www.nutriset.fr/en/product-range/prevention-of-malnutrition.html.

5. Tables

Table 1: Baseline Respondent, Household, and Maternal Characteristics

	Variable	Definition	Mean/ Count	Std Dev/ Percent	Min	Max
Respondent	Head of Household	*= 1 if respondent is head of household (=0 if primary caregiver)*	xxx	xx.xx		
	Age	*Age in years*	xx.xx	xx.xx	xx	xx
	Education	*Years of Education*				
Household	Children U5	*Number of household members who are children under five years*				
	Percent Children U5	*Percent of children under five years in household*				
	HFIA Score	*Household Food Insecurity Access Score*				
	Asset Index	*Proxy for socioeconomic status*				
	Per Capita Total Expenditures	*Per capita daily total expenditures (4th Quarter 2011 USD)*				
	Percent Food Expenditures	*Percent of total household expenditures on food*				
	Discount Rate	*Relative measure of time discounting*				
	Risk Behavior	*Relative measure of willingness to take a risk*				

Child	First Child	*= 1 if iLiNS child's mother has no other children*
	LNS	*= 0 if child in delayed intervention group =1 otherwise*
	Male	*=1 if iLiNS child is male (= 0 if iLiNS child is female)*
	Height-for-Age	*Height-for-Age Z-score*
Mother	Age	*Age in years*
	Education	*Years of Education*
	Height	*Height in meters*
	BMI	*Body mass index*

N=xxx

Significance codes: ***(p<.01), **(p<.05), *(p<.1) indicate difference in means between LNS and the delayed intervention groups.

Table 2: Average WTP for a Week's Supply of LNS and LP at Baseline

Product	N	Mean† (Std Error)	Std Dev	Min	Max*	Zero WTP/ Difference
LNS	xxx	x.xx (x.xx)	x.xx	x.xx	x.xx	xx (x.x%)
LP						
LNS - LP						

†In 4th Quarter 2011 US Dollars
*Observations>4 SD above the mean were omitted as outliers.
Significance codes: ***(p<.01), **(p<.05), *(p<.1) indicate different mean WTP across products.

Table 3: Average WTP for a Week's Supply of LNS and Likuni Phala at Baseline by Treatment Group and Respondent

Product		N	Mean	Std Error
LNS	Overall	xxx	x.xx	x.xx
	LNS			
	Non-LNS			
	iLiNS Woman			
	Head of Household			
Likuni Phala	Overall			
	LNS			
	Non-LNS			
	iLiNS Woman			
	Head of Household			
LNS - LP	Overall			
	LNS			
	Non-LNS			
	iLiNS Woman			
	Head of Household			

†In 4th Quarter 2011 US Dollars.
For treatment groups: significant codes ***(p<.01), **(p<.05), *(p<.1) indicate mean WTP for LNS group different than delayed intervention group for same product.
For respondents: significant codes ***(p<.01), **(p<.05), *(p<.1) indicate mean WTP for iLiNS women different than head of households for same product.

Table 4: Average Long-Term WTP for LNS and Likunia Phala at Baseline

Product	N	Mean† (Std Error)	Std Dev	Min	Max*	Zero WTP/ Difference
LNS	xxx	x.xx (x.xx)	x.xx	x.xx	x.xx	xx (x.x%)
Likuni Phala						
LNS - LP						

†In 4th Quarter 2011 US Dollars.
*Observations > 4 SD above the mean were omitted as outliers.
Significance codes: ***(p<.01), **(p<.05), *(p<.1) indicate different mean WTP across products.

Table 5: Average Long-Term WTP for LNS and Likuni Phala at Baseline by Treatment Group and Respondent

Product		N	Mean	Std Error
LNS	Overall	xxx	x.xx	x.xx
	LNS			
	Non-LNS			
	iLiNS Woman			
	Head of Household			
Likuni Phala	Overall			
	LNS			
	Non-LNS			
	iLiNS Woman			
	Head of Household			
LNS - LP	Overall			
	LNS			
	Non-LNS			
	iLiNS Woman			
	Head of Household			

†In 4th Quarter 2011 US Dollars.
For treatment groups: significant codes ***(p<.01), **(p<.05), *(p<.1) indicate mean WTP for LNS group different than delayed intervention group for same product.
For respondents: significant codes ***(p<.01), **(p<.05), *(p<.1) indicate mean WTP for iLiNS women different than head of households for same product.

Table 6: Regression Results - Baseline WTP for a Week's Supply

		Coefficient (Robust Standard Error)		
	Variable	LNS	LP	Difference
Respondent	Head of Household (0/1)	x.xxx (x.xxx)	x.xxx (x.xxx)	x.xxx (x.xxx)
	Age (yrs)			
	Education (yrs)			
Household	Children U5			
	Percent Children U5			
	HFIA Score			
	Asset Index			
	Per Capita Total Expenditures (USD)			
	Percent Food Expenditures			

	Discount Rate			
	Risk Behavior			
Child	First Child (0/1)			
	LNS (0/1)			
	Male (0/1)			
	Height-for-Age			
Mother	Age (yrs)			
	Education (yrs)			
	Height (meters)			
	BMI			
	Constant			
	N	xxx	xxx	xxx
	R^2	x.xxx	x.xxx	x.xxx

Significance codes: ***(p<.01), **(p<.05), *(p<.1)
Note: Controls for month and year of enumeration, enumerator, and language of enumeration were also included in the model (unreported).

Table 7: Regression Results - Long-Term Baseline WTP

		Coefficient (Robust Standard Error)		
	Variable	LNS	LP	Difference
Respondent	Head of Household (0/1)	x.xxx (x.xxx)	x.xxx (x.xxx)	x.xxx (x.xxx)
	Age (yrs)			
	Education (yrs)			
Household	Children U5			
	Percent Children U5			
	HFIA Score			
	Asset Index			
	Per Capita Total Expenditures (USD)			

	Percent Food Expenditures
	Discount Rate
	Risk Behavior
Child	First Child (0/1)
	LNS (0/1)
	Male (0/1)
	Height-for-Age
Mother	Age (yrs)
	Education (yrs)
	Height (meters)
	BMI
	Constant

N	xxx	xxx	xxx
R^2	x.xxx	x.xxx	x.xxx

Significance codes: ***(p<.01), **(p<.05), *(p<.1)

Note: Controls for month and year of enumeration, enumerator, and language of enumeration were also included in the model (unreported).

iLiNS-DOSE

Sustained acceptability of LNS (lipid-based nutrient supplements) for infants in Malawi

Statistical analysis plan (SAP)

Appendix 07: Long-term acceptability of LNS (lipid-based nutrient supplements) for infants in Malawi (added on 28.05.2014)

Contents

1. Overview

The effectiveness of any complementary feeding intervention depends on both its biological efficacy and its sustained acceptability among the target population. In earlier research, LNS products have been acceptable during brief exposures (often 2 weeks) but sustained acceptability has not been investigated. The statistical analysis plan (SAP) below describes an analysis intended to look at sustained acceptability of LNS as part of the Malawi iLiNS-DOSE trial.

2. Background and objectives

In the iLiNS-DOSE trial, we did not observe any statistically significant growth promotion among children whose diets were supplemented with lipid based nutrient supplements (LNS) between 6 to 18 months. Based on previous literature, other sub-studies with iLiNS DOSE trial (Hemsworth, unpublished PhD Thesis) and trial participants' responses during in-depth interviews, we assume that this minimal growth effect could partially be attributed to the way families were able to integrate the products in the diets of the children.

At enrolment and repeatedly during the follow-up, the study team made the following recommendations about to the use of LNS:

1. the mothers should be giving LNS to the study child every day throughout the 12-month intervention period,
2. LNS should be given only to the study child, and
3. LNS should be mixed with a small amount of food before feeding.

Along with these recommendations, the team provided the guardians with some generic advice on nutrition and breastfeeding.

In the current analysis we will study the sustained acceptability of the LNS products by examining the guardians' adherence to the above described recommendations on LNS utilization. We do not aim to calculate an aggregated indicator for adherence but will present the analysis by three variables separately. Of these, adherence percentage (the proportion of days when the participant was offered LNS) and deviation from the recommendation of daily provision of LNS will indicate adherence to the recommendation of giving LNS every day to the study child. Supplement sharing and deviation from the recommended feeding mode will indicate adherence to the two other recommendations. Where appropriate, we will examine these components of adherence by intervention group or as a function of time.

Besides the immediate adherence variables, we will describe underlying acceptability factors that indirectly affect LNS use in the study community. Themes included in this analysis include e.g. the perceived ease of use of LNS products, their assumed impacts on child growth and health, and their social acceptability. A concept map that describes the relationships between acceptability, adherence, and eventual LNS consumption is shown in appendix 1.

According to our concept map, the guardian reported consumption of LNS is determined by three factors during the process: participant retention in the study (impaired by participants' withdrawal of consent), the study team's success in the timely delivery of the supplements (impaired mostly by the participants permanently or temporarily moving outside the study area), and the guardians' adherence to daily supplement feeding recommendation To separate the effect of the sustained acceptability and the other two determinants of the LNS consumption, the adherence variables will not include any data from days when LNS was not available to the study participants. However, whilst the study discontinuation or a move to other areas was mostly not affected by the acceptability of LNS, we will report also the delivery success, to get a more comprehensive picture of supplement consumption in the study population.

3. Data

Data were collected with different methods:

1. Data collectors visited the homes of participants weekly throughout the intervention in order to collect information pertaining to the previous week. Mothers were asked to daily record LNS use in a picture chart, which was used as a memory aid during brief weekly consumption surveys (form 18).
2. After six and twelve months in trial (infant age 12 and 18 months, respectively), mothers responded to a detailed knowledge, attitudes and practices (KAP) survey. As part of this, they were asked to report the way they had utilized LNS during the preceding week (form 15 b and c).
3. At four and ten months into the trial (infant age 10 and 16 months, respectively), there were structured observations conducted at the participants' homes.
4. At about one and seven months into the trial (infant age 7 and 13 months, respectively), 30 guardians underwent in-depth interviews about the experience on LNS use.

The first three data sets will be used for the current quantitative analyses, whereas the fourth dataset will be used to supplement the quantitative section with qualitative analyses

4. General approach to data

A major part of the analysis will be driven by pre-defined hypotheses. Additionally, there will be a number of exploratory analyses on mother, child and context related factors that might predict sustained acceptability.

The delayed intervention group (control group) will not be included in the analysis, because children in this group never received any LNS.

The remaining study participants initially formed five separate groups, based on the type of LNS (milk-containing or milk-free) and the daily dose of it (10 grams with milk, 20 or 40 grams / day with or without milk). In the analysis, we will first test if there are any statistically significant differences

in adherence between guardians whose children received either milk-containing or milk-free LNS products. We will do this initial analysis using only our main outcome variable, i.e. adherence percentage. If there are no statistically significant differences in adherence between the two groups, we will collapse the groups and place the participants into three aggregated groups. These groups will be defined by the dose of LNS provided to the participant (10g, 20g, or 40g / day).

By participant we mean the mother- child dyad that was enrolled to the study. In some cases, the mother may have been subsequently replaced by another guardian.

We will use Stata software (version 12, StataCorp, Texas) for the statistical analysis. For qualitative data, we will useATLAS.ti software (version 5.7.1, Cincom Systems Inc) that enables analysis of acceptability related emerging themes and subsequent thematic analysis.

5. Hypotheses to be tested

Our assumption is that both the milk-containing and milk-free LNS products were in general well accepted, but that there may have been a decrease in adherence over time. We further hypothesize that provision of larger doses of LNS to the household did not increase the actual ration consumed by the study child, due to more frequent sharing of the supplement to others than the study participant.

More specifically, the hypotheses to be tested are:

1. The mean adherence percentage during a 12-month LNS supplementation period is not lower among infants provided with milk-free LNS than among infants provided with milk-containing LNS.

2. The mean adherence percentage over the entire LNS supplementation period is not lower among infants provided with the 10 g / day dose than among infants provided with 20 g / day or 40 g / day of LNS.

3. The mean adherence percentagr decreases as a function of time during a 12-month LNS supplementation period.

4. The proportion of mothers reporting deviating from the recommendation of daily provision of LNS to their infants is higher after 12 than after 6 months of supplementation.

5. The proportion of mothers reporting supplement sharing is higher after 12 than after 6 months of supplementation.

6. The proportion of mothers reporting supplement sharing is higher among infants provided with the 40 g / day dose than among infants provided with 20 g / day or 10 g / day of LNS.

6. The proportion of mothers deviating from the recommended feeding mode (LNS to be mixed with food) is higher after 12 than after 6 months of supplementation.

6. Data cleaning and breaking the code

This will be done abiding to the procedures developed for the main trial.

7. Definition of the outcome variables

We will primarily determine an overall adherence percentage for each participant by calculating the proportion of days when the supplement was reportedly used by the participating infant.

- The formula to be used is: Adherence percentage = number of days during the follow-up period when the supplement was reportedly used (answer alternative 1)/ total number of days with data on LNS consumption during the follow-up period (answer alternative 0 or 1)* 100%
- Source of data: question 2.4 (HomLNSconsumed) in form 18
- For the testing of hypothesis 1 and 2, all available data for each participant will be combined into aggregated value that represents the mean overall adherence percentage during the entire follow-up. We chose this approach, instead of using repeated measurement analysis because of the small observed variation in the weekly adherence percentage. For the participants in five LNS intervention groups we have data for about 67,200 weekly consumption and less than 10% of these report having consumed LNS on less than 100% of the days during last week.
- For the testing of hypothesis 3, we will calculate each participant's weekly adherence percentage separately for each week in follow-up (visit v01-visit v52).

Secondarily, we will use data from the KAP interviews to determine weekly adherence percentage at two specific time points. At 6 and 12 months into the trial, the participants were asked about the number of days when the study child was fed with LNS on the preceding week (7 days). Maximum number is 7, minimum 0.

- Source of data: question 3.1 in form 15b and question 5.1 in form 15c (KapEatSevenDay)
- For alternative ways of testing the hypotheses 2 and 3, we will calculate each participant's weekly adherence percentage separately for the two time points when KAP interviews were done.

For the testing of hypothesis 4, we will determine deviation from the recommendation of daily provision of LNS for each participant at two time points and by intervention group, by defining the participants who reported not feeding LNS on every day during the preceding 7 days (answer options 0-6)

- Data with form 15b was collected at 6 months after enrolment
- Data with form 15c was collected at 12 months after enrolment

- Source of data: question 3.1 in form 15b and question 5.1 in form 15c (KapEatSevenDay)

For the testing of hypotheses 5 and 6, we will determine supplement sharing for each participant at two time points and by intervention group, by defining the participants who reported sharing (answer options 2-8 or 66) during any of the days during the week preceding the interview

- Data with form 15b was collected at 6 months after enrolment
- Data with form 15c was collected at 12 months after enrolment
- Source of data: form 15b, question 3.5 and 15c, question 5.5, (KapSharing).

For the testing of hypothesis 7, we will determine deviation from the recommended feeding mode for each participant at two time points, by determining the participants who reported having provided LNS in any other mode than mixed with food during the week preceding the interview.

- Data with form 15b was collected at 6 months after enrolment
- Data with form 15c was collected at 12 months after enrolment
- Source of data: form 15bquestion 3.12 and 15c question 5.12, (KapMixFood).

8. Safety outcome

Not applicable.

9. Basis for analysis

The analysis will be based on the principle of modified intention-to-treat (see main analysis plan for modification). Thus we will use the randomly assigned group allocation as the independent variable. All the analysis will exclude children who did not get supplement during the interventions (control group). Children who have form 18 data collected 8 times or more will be included in this analysis. This means that about 7% of the participants will be excluded.

All participants who have data for form 15 will be included in the analysis. Because of the rather high number of missing values for variables from form 15 (20%-36% depending on the outcome) we will use multiple imputation to impute missing values.

10. Time points for the analyses

Form 18 data were collected through-out the study and length of follow-up will be calculated as part of the daily consumption variable. Form 15 data will be analyzed at two time points (after child had been six and twelve months in trial, infant age 12 and 18 months, respectively.

11. Imputation of missing values for selected outcome variables.

Values for the outcome variables derived from form 15 KAP interviews (number of days when LNS was consumed last weeks, recommended daily provision of LNS, feeding mode of LNS and sharing of LNS) are missing for 22%-36% participants (neither 12 nor 18 month data available). We will use multiple imputation to impute these variables and compare the results to those done without multiple imputation.

Multiple imputation model will be based on 10 – 15 variables describing the intervention group as well as selected maternal, child and household related factors. Maternal factors will include e.g. mother's age, literacy, marital status, and number of children. Child-related variables will include age at data collection and weight for height and height for age at enrollment and at 18 months of age. Household related factors will include home village, household asset score, and food security score.

12. Presentation of the study findings and hypotheses testing

Since varying numbers of data are available depending on the time point, sample sizes by group will be reported for each time point. If a dependent variable is missing for more than 10% of infants (with denominator being total records available for the time point) we will report the number of observations used per specific outcome analysis.

Enrolment to the study was the same as for the main study and we will refer for success of enrolment presented in the main article.

Baseline information

We will tabulate selected summary characteristics at enrollment by intervention arms, as indicated in table 1. We will test hypotheses about a difference between the three intervention groups using chi-square for categorical variables and ANOVA and Kruskal-Wallis test for continuous variables with normal and non-normal distribution, respectively. However, p-values obtained from these tests will not be shown in the eventual articles.

Comparison of adherence between participants receiving either milk-containing or milk-free LNS

We will test hypothesis 1 on a differences in the adherence percentage between participants receiving milk-containing and milk-free LNS using the adherence percentage data from form 18. This analysis will exclude participants who received 10g / day milk-containing LNS. Regression model of the below form will be estimated:

$$\hat{y}_i = b_0 + b_1 D_i + b_2 S_i + b_3 (D_i \times S_i)$$

where D = 0 or 1 for the 20g and 40g group, respectively, and S = 0 or 1 for the milk and milk-free LNS group, respectively. The model will be shown as illustrated in Table 2.

Coefficient b_2 will indicate whether the 20g milk-LNS group differs from 20g no-milk-LNS group and linear combination of coefficients $b_2 + b_3$ (Stata's *lincom* command) will tell us whether the 40g

milk-LNS group differs from 40g no-milk-LNS group. If P<0.05 for coefficient b_2 or $b_2 + b_3$ we will reject the null hypothesis of no differences between groups and use separate groups for 20g and 40gmilk and no-milk LNS groups in the following analyses. If P>0.05 for coefficient b_2 or $b_2 + b_3$ we will combine milk and no-milk LNS groups in the following analyses.

Comparison of adherence between participants receiving 10 g, 20g or 40 g / day of LNS

The distribution of individual adherence percentages over the entire follow-up will be shown by study group (10g, 20g, 40g intervention), using a box and whiskers plot as shown in figure 1. Both mean and median adherence percentages will be shown in the Box-and-whiskers plot.

The formula used for determining the mean adherence percentage by study group is:
The sum of the individual adherence percentage / the number of participants with data.

For testing hypothesis 2 (about a difference in adherence between participants receiving different doses of LNS) we will first construct a regression model, with the mean adherence percentage as the dependent variable and the intervention group as independent variable (table 3 unadjusted model).

We will then create additional bivariate regression models to find out which other study variables predict the adherence percentage (table 3, unadjusted model). We will present the constant (95% CI) of the regression model and the deviation (95% CI) from the constant for each variable. For categorical independent variables the constant will be the mean value of the reference group and the regression coefficient presents the difference between the groups. For continuous variables the constant will present the predicted outcome when the predictor is 0 and the coefficient presents the difference in the predicted outcome per one unit increase in the exposure variable (Cheung, 2013).

Finally, we will construct an adjusted regression model for the adherence percentage (table 3, adjusted model). The covariates to be included in the analysis will be those that had an association (P-value < 0.1) in the bivariate analysis. The coefficients present the change in the outcome variable per one unit change in the covariate when all the other covariates are held constant (Cheung, 2013)

The change of adherence percentage as function of time

The mean adherence percentage by study group (10, 20, or 40 g LNS / day, and all groups combined, i.e. 4 lines in total) will be plotted as a function of study week, in a line chart and using a 6-week average for each time point (figure 2).

For hypothesis testing (hypothesis 3 on a decrease in adherence percentage over time), we will calculate the rate of change in the weekly adherence percentage by study group. We will create a regression model and calculate a slope for weekly adherence percentage for each participant having form 18 collected 8 times or more. We will store the individual regression coefficients in a new variable and calculate the mean change of weekly adherence percentage by study group. A hypothesis about a decreasing weekly adherence percentage over time will be tested with a least squares regression with individual slopes for change of weekly adherence percentage as dependent variable and study group as independent variable (Cheung, 2013). Null hypothesis of no change in adherence percentage over time will be rejected if P<0.05 for any of the group coefficients.

For the alternative definition of weekly adherence percentage, we will use data collected from the KAP interviews at two time points (after 6 and 12 months of LNS provision). At both time points, we will first calculate the mean weekly adherence percentage and the proportion of participants who report deviation from the recommendation of daily LNS provision. The formula to be used in the calculation of the mean weekly adherence percentage proportion will be the following:

- The sum of individual weekly adherence percentages / total number of guardians providing data on this question * 100%.

The formula to be used in the calculation of the proportion of deviating participants will be the following:

- The number of guardians reporting that their infant consumed LNS on anything else than 7 days during the preceding week (answer alternative 0-6)/ total number of guardians providing data on this question * 100%.

The group summaries will be shown in Table 4. For hypothesis testing (alternative test for hypothesis 3 on a decrease in adherence percentage over time), we will create a mixed model with random intercept and random slope for time and interaction terms between the study group and time of measurement. We will reject the null hypothesis of no change in adherence percentage over time if coefficients for intervention group and time interaction $P<0.05$. Results will be shown in the text.

For testing hypothesis 4 (deviation from daily LNS provision) we will create a logistic regression model with deviating from daily LNS provision as a dependent variable and visit as an independent variable (intervention groups combined). With this model we will test the null hypothesis of no difference in deviation from daily LNS provision between two time points and the null hypothesis will be rejected if $P<0.05$. Results of this test will be shown in the text. Next we will test the change in deviation from daily LNS provision over time by intervention group. We will create a mixed model with random intercept and random slope for time and interaction terms between the study group and time of measurement. We will reject the null hypothesis of no change in daily LNS provision over time if $P<0.05$.

<u>Supplement sharing</u>

At both time points, we will calculate the proportion of participants who report supplement sharing (table 4). The formula to be used in the calculation of the proportion of deviating participants will be the following:

- number of guardians reporting that they gave the LNS supplement to someone else than the study child (answer alternatives 2 to 8, 66)/ total number of guardians providing data on this question * 100%.

For hypothesis testing (hypothesis 5 on more frequent sharing after 12 than after 6 months of supplementation) we will first create a logistic regression model with sharing as a dependent variable and visit as an independent variable (intervention groups combined). With this model we will test the null hypothesis of no difference in sharing between two time points and the null hypothesis will be rejected if $P<0.05$. Results of this test will be shown in the text. Next we will test the change in sharing over time by intervention group. We will create a mixed model with random intercept and

random slope for time and interaction terms between groups and time of measurement. We will reject the null hypothesis of no change in deviation from the recommendation of daily LNS provision over time if $P<0.05$ (table 5).

We will further present the proportion of participants reporting LNS sharing at the two time points, stratified by intervention group (table 6). Hypothesis 6 (on a difference in sharing between the study groups) will be tested with the same logistic mixed model as hypothesis 5. We will calculate difference in odds ratios of sharing at the two time points separately between 10g or 20g and 40g LNS groups. The null hypothesis of no differences between 10g or 20g and 40g LNS groups will be rejected if $P<0.05$.

Deviation from the recommended feeding mode

At both time points, we will calculate the proportion of participants who report deviation from the recommended feeding mode (Table 4). The formula to be used in the calculation of the proportion of deviating participants will be the following:

- number of guardians reporting that they gave the LNS supplement in some other mode than mixed with food (answer alternatives 1,2,3)/ total number of guardians providing data on this question * 100%.

For hypothesis testing (hypothesis 7 on a change in deviations from the recommendation over time) we will create a logistic regression model with deviation from the recommendation as a dependent variable and visit as an independent variable (intervention groups combined). With this model we will test the null hypothesis of no difference in deviation from recommended feeding mode between two time points and the null hypothesis will be rejected if $P<0.05$. Results of this test will be shown in the text. Next we will test the change in sharing over time by intervention group. We will create a mixed model with random intercept and random slope for time and interaction terms between groups and time of measurement. We will reject the null hypothesis of no change in deviation from the recommended feeding mode over time if $P<0.05$ (Table 5).

Exploratory analysis

We will include in table 4 data on the mean number of feedings on the preceding day, and the proportion of guardians reporting that feeding LNS was difficult. For these data, we will not perform any statistical tests on possible differences between the values by study group or time point.

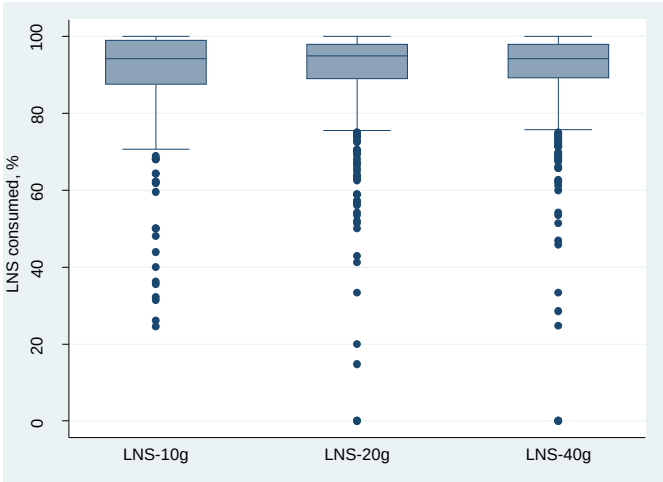
Observations

In the text, we will count the proportion of participants who were observed consuming LNS and tabulate it by LNS intervention group. And for those who were seen to receive LNS, we will count the mean number of LNS servings given per day and tabulate that by LNS Intervention.

13. Design of figures and tables

The below figures and tables will be prepared based on the analysis:

Figure 1: Box-whisker plot of adherence percentage among participants receiving 10g, 20g, 40g LNS supplementation per day.



* mean and median for LNS-10g, LNS-20g and LNS-40g are, xxx, xxx, xxx, respectively.

Figure 2: The mean reported weekly adherence percentage among participants receiving 10g, 20g, 40g LNS supplementation per day.

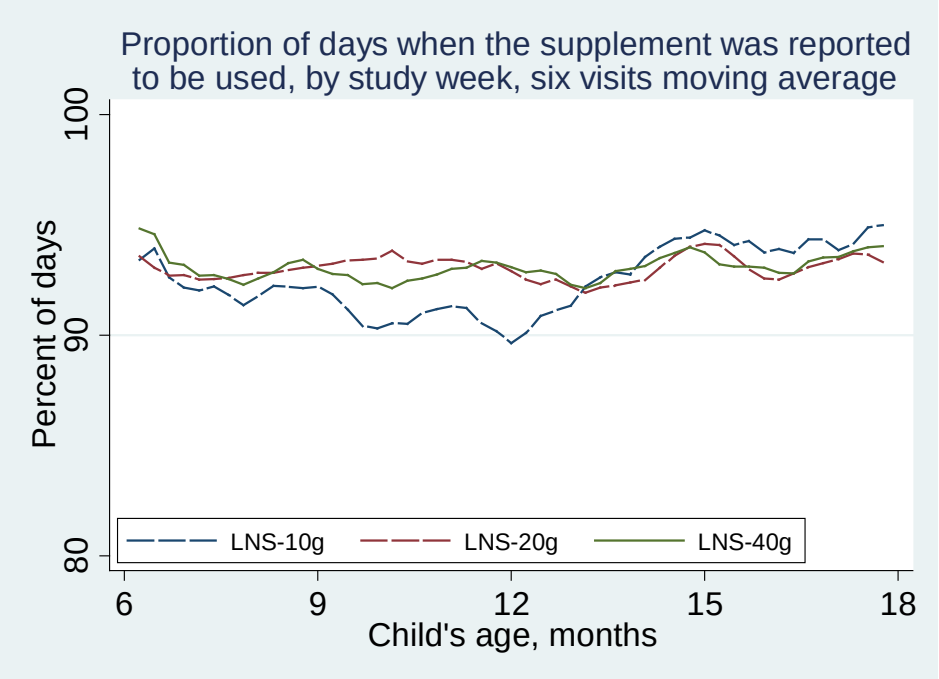


Table 1: Participant and household characteristics at enrollment, by study group.

Characteristic	LNS-10 group	LNS-20 group	LNS-40 group	P-value (test chi-square or anova/Kruskal-Wallis)
Child				
Number of participants	xxx	xxx	xxx	
Age, months, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Length for age, adjusted for sex, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Weight for height, adjusted for sex, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Proportion of girls	xx %	xx %	xx %	0.xxx
Prevalence of moderate and severe stunting at 6 months of age	xx %	xx %	xx %	0.xxx
Mother				
Age years, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Schooling years, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Proportion of married	xx %	xx %	xx %	0.xxx
Proportion of women with only one child	xx %	xx %	xx %	0.xxx
Household				
Household asset index, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Number of u-5 children in family, mean (SD)	xx (xx)	xx (xx)	xx (xx)	0.xxx
Head of household is respondent's husband	xx %	xx %	xx %	0.xxx
Proportion enrolled from Mangochi	xx	xx	xx	0.xxx

Table 2 The association between the milk content in LNS and the guardian reported LNS adherence percentage, by the dose of LNS: results from a regression analysis.

	Model: $\hat{y}_i = b_0 + b_1 D_i + b_2 S_i + b_3 (D_i \times S_i)$		
Variable	**Coefficient**	**90 % CI**	**P-value**
Constant (b_0)	xx.xx	xx.xx to xx.xx	x.xxx
Difference in the mean adherence percentage between participants receiving 20 g / day milk-free LNS and participants receiving 20 g / day milk-containing LNS (b_2)	xx.xx	xx.xx to xx.xx	x.xxx
Difference in the mean adherence percentage between participants receiving 40 g / day milk-free LNS and participants receiving 40 g / day milk-containing LNS ($b_2 + b_3$)	xx.xx	xx.xx to xx.xx	x.xxx

Table 3: The association between LNS dose, selected participant and household characteristics and LNS adherence percentage.

		Unadjusted analysis		**Adjusted analysis**	
Predictor variable	**n**	**Coefficient (95% CI)**	**P-value**	**Coefficient (95% CI)**	**P-value**
Daily LNS dose					
10 g (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
20 g	xxx	x.x (x.x to x.x)	0.xxx		
40 g	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Time in study, weeks	xxx	x.x (x.x to x.x)	0.xxx		
Child sex					
Boy (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
Girl	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Child LAZ at enrolment, Z-score	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Child WHZ at enrolment, Z-score	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Maternal age, years	xxx	x.x (x.x to x.x)	0.xxx		
Number of children					
1 (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
2	xxx	x.x (x.x to x.x)	0.xxx		
3+	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Maternal education, years	xxx	x.x (x.x to x.x)	0.xxx		
Maternal ability to read					
None (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
With difficulty	xxx	x.x (x.x to x.x)	0.xxx		
Fluently	xxx	x.x (x.x to x.x)	0.xxx		
Maternal ability to write					
None (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
With difficulty	xxx	x.x (x.x to x.x)	0.xxx		

Fluently	xxx	x.x (x.x to x.x)	0.xxx		
Maternal marital status					
Married (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
Single	xxx	x.x (x.x to x.x)	0.xxx		
Other	xxx	x.x (x.x to x.x)	0.xxx		
Maternal tribe					
Chewa (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
Yao	xxx	x.x (x.x to x.x)	0.xxx		
Other	xxx	x.x (x.x to x.x)	0.xxx		
Head of household					
Head of household is respondent's husband (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
Head of household is not respondent's husband	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Number of children under-5	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Household asset index, index-score	xxx	x.x (x.x to x.x)	0.xxx		
Constant		xx.x (x.x to x.x)	0.xxx		
Household food security, index-score	xxx	x.x (x.x to x.x)	0.xxx		
Site of residence					
Mangochi (Constant)	xxx	xx.x (x.x to x.x)	0.xxx		
Other	xxx	x.x (x.x to x.x)	0.xxx		

Table 4: Participant adherence to feeding recommendations and experience from LNS use; results from KAP interviews at two time points.

	Duration of LNS provision	
	6 months	12 months
Mean (SD) weekly adherence percentage (on 7 days in the preceding week)	xx.x (xx.x)	xx.x (xx.x)
Proportion of guardians reporting deviation from the recommendation of daily LNS provision during the preceding week	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)
Proportion of guardians reporting LNS sharing during the preceding week	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)
Proportion of guardians reporting deviation from the recommended feeding mode (mix with food) during the preceding week	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)
Mean (SD) number of LNS feedings on the preceding day	xx.x (xx.x)	xx.x (xx.x)
Proportion of guardians reporting that feeding LNS was difficult	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)

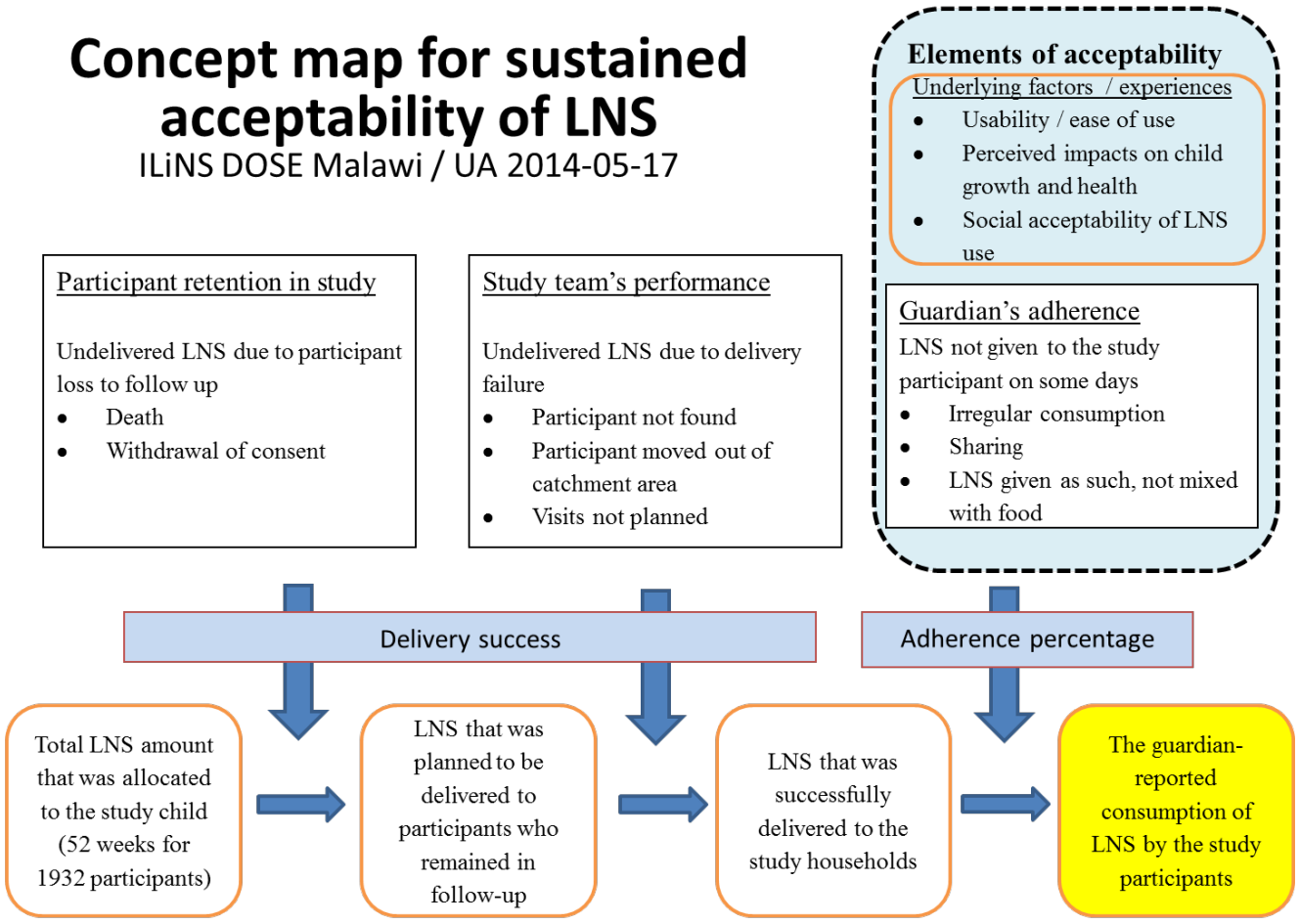
Table 5. The association between daily dose of supplemented LNS and change in adherences to the recommendations on LNS use between two time points. Results from a logistic mixed model

	Model: $\hat{y}_i = b_0 + b_1 time_{ij} + b_2 LNS10g_i + b_3 LNS20g_i + b_4(LNS10g_i \times time_{ij}) + b_5(LNS20g_i \times time_{ij})$		
Outcome variable	**Daily LNS dose**	**OR (95% CI)**	**P-value**
Change in odds ratio over time for guardians reporting deviation from the recommendation of daily LNS provision during the preceding week	10 g (b_4 x time)	xx.x (x.x to x.x)	0.xxx
	20 g (b_5 x time)	xx.x (x.x to x.x)	0.xxx
	40 g (b_1 x time)	xx.x (x.x to x.x)	0.xxx
Change in odds ratio over time for guardians reporting LNS sharing during the preceding week	10 g (b_4 x time)	xx.x (x.x to x.x)	0.xxx
	20 g (b_5 x time)	xx.x (x.x to x.x)	0.xxx
	40 g (b_1 x time)	xx.x (x.x to x.x)	0.xxx
Change in odds ratio over time for guardians reporting deviation from the recommended feeding mode (mix with food) during the preceding week	10 g (b_4 x time)	xx.x (x.x to x.x)	0.xxx
	20 g (b_5 x time)	xx.x (x.x to x.x)	0.xxx
	40 g (b_1 x time)	xx.x (x.x to x.x)	0.xxx

Table 6: The association between daily dose of supplemented LNS, time in study, and supplement sharing. Results from a logistic mixed model presented in Table 5.

	Daily LNS dose		
	10g	20g	40g
Proportion of guardians reporting LNS sharing during the preceding week, at 6 months	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)
Difference in OR between 10g or 20g dose and 40g dose	xx.x (xx.x to xx.x)	xx.x (xx.x to xx.x)	Ref.
P-value	xxx.x	xxx.x	
Proportion of guardians reporting LNS sharing during the preceding week, at 12 months	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)	xxx/xxx (xx.x%)
Difference in OR between 10g or 20g dose and 40g dose	xx.x (xx.x to xx.x)	xx.x (xx.x to xx.x)	Ref.
P-value	xxx.x	xxx.x	

Appendix 1:


Concept map for sustained acceptability of LNS
ILiNS DOSE Malawi / UA 2014-05-17
Elements of acceptability
Underlying factors / experiences
Usability / ease of use
Perceived impacts on child growth and health
Social acceptability of LNS use
Participant retention in study
Undelivered LNS due to participant loss to follow up
Death
Withdrawal of consent
Study team's performance
Undelivered LNS due to delivery failure
Participant not found
Participant moved out of catchment area
Visits not planned
Guardian's adherence
LNS not given to the study participant on some days
Irregular consumption
Sharing
LNS given as such, not mixed with food
Delivery success
Adherence percentage
Total LNS amount that was allocated to the study child (52 weeks for 1932 participants)
LNS that was planned to be delivered to participants who remained in follow-up
LNS that was successfully delivered to the study households
The guardian-reported consumption of LNS by the study participants

Prevention of Linear Growth Faltering in Infants and Young Children with Lipid-based Nutrient Supplements (iLiNS-DOSE)

Statistical analysis plan

Appendix 08: The effect of Lipid-based Nutrient Supplement (LNS) provision on child morbidity (Added on 28.05.2014)

Contents

1 Version history

Version number	Version date	Prepared by	Description of completed editions
01.0	27/05/2014	Jaden Bendabenda	Original document

2 Introduction

Although nutritional supplementation is one of the recommended ways to combat child undernutrition, emerging evidence suggests that the supplements may be harmful in some instances. While a beneficial effect of nutrient supplementation on morbidity has been shown in some studies (Sharieff et al. 2006), big studies have shown that supplementation with multiple micronutrient powders or iron tablets may be associated with increased risk of common childhood infections (Sazawal et al, 2006, Soofi et al, 2013). Lipid-based nutrient supplement (LNS) is a product being tested for its potential in prevention of under nutrition and improving child growth outcomes in a randomised controlled trial in a rural Malawian population. It may be important to provide evidence of safety of LNS provision in settings where childhood infections are common. Previous studies on LNS have reported that the product appears to be safe but the evidence is still not conclusive. Such studies had either relatively short duration, i.e. 6 months or less, or did not have sufficient power because of small sample size (Dewey & Baldiviez, 2012). This study aims at assessing the safety of LNS when provided to healthy infants and young children for a period of 12 months. We will also explore any beneficial effects of LNS provision on morbidity, as an additional outcome.

3 Study objective

The main aim of this study is to assess the safety outcomes in children participating in the iLiNS-DOSE study in rural communities in Malawi. The outcomes are categorised as serious adverse events (SAE), non-scheduled visits, self-reported morbidity symptoms and self-reported disease episodes.

3.1 Specific objectives

The specific objective of the study is to compare the following outcomes between the study groups after 12 months of intervention:

i. Risk of SAEs (hospitalizations and deaths).
ii. Incidence of non-scheduled visits made by the participants to health facilities due to malaria or any other illnesses.
iii. Longitudinal prevalence of self-reported common childhood morbidity symptoms.
iv. Incidence of common childhood infectious disease episodes.

4 Hypotheses to be tested

a) Risk of SAEs (deaths or hospitalizations) is not markedly higher in children receiving LNS than in children who receive no dietary intervention.
b) Incidence of non-scheduled visits made to health facilities due to malaria or any other illnesses is not markedly higher in children receiving LNS than in children who receive no dietary intervention.
c) Longitudinal prevalences of self-reported morbidity symptoms are not markedly higher in children who receive LNS than in children who receive no dietary intervention at the same age.
d) Incidences of self-reported episodes of common childhood infectious disease episodes (malaria, acute respiratory illnesses and gastroenteritis) are not markedly higher in children who receive LNS than in children who receive no dietary intervention at the same age.

5 Definitions of end-points

Serious adverse events (SAE): Defined according to United States Department of Health and Human Services Office for Human Research Protections (www.fda.gov). *The data will be extracted from form 29, Q 3.2.*

Non-scheduled visits: Non-scheduled visits are a composite of morbidity, access to health care and health seeking behaviour. The total number of visits to a health facility made by the participant during the study period, as well as clinical diagnoses made by health workers at the

non-scheduled visits will be analysed. *The data will be extracted from form 26 (total) and form 26, Q 4.1.*

Self-reported symptoms of any illness for each day were collected at a weekly home visit. These data will be extracted from form 18 where the following symptoms were recorded: *low appetite (Q 2.6), diarrhoea (Q 2.7), bloody stools (Q 2.8), mucoid stools (Q 2.8), vomiting (Q 2.9), fever (Q 2.10), cough (Q 2.11), rapid breathing (Q 2.12), difficult breathing (Q 2.13) and nasal discharge (Q 2.14).*

Diagnoses of malaria, gastroenteritis and acute respiratory infections (ARI) were derived from presence of a combination of self-reported symptoms on one or more days. To ensure that the diagnoses are mutually exclusive, a diagnosis algorithm was created whereby any diarrhea episode (≥3 loose stools in 24 hours) was categorized as gastroenteritis in the presence or absence of other symptoms. If diarrhea was absent but there was presence of any respiratory symptoms (cough, rapid or difficult breathing and nasal discharge) with or without fever, a diagnosis of ARI was made. Fever episodes in the absence of diarrhea and respiratory symptoms, with or without other symptoms, were categorized as malaria. The rest of the symptoms in the absence of diarrhea, respiratory symptoms and fever were categorized as "others". For all diseases, an episode was defined as the period starting from the day the child had symptoms when preceded by at least 2 days of no symptoms or no data. The episode ended on the last day the child had symptoms which was then followed by at least 2 symptom-free days.

6 Basis of analysis: Intention to treat and per protocol analysis

A total of 1932 six-month-old children were randomized into five intervention groups to receive the following daily doses: 10 g milk-LNS, 20 g milk-LNS, 20 g non-milk-LNS, 40 g milk-LNS, 40 g non-milk-LNS and a control (no dietary intervention) group. In the first analysis we will compare the outcomes among all six groups, with the aim of testing the difference between the milk and milk-free groups. We will test the hypothesis of no difference in the longitudinal prevalence of any illness using least squares regression. The regression model will be estimated as follows:

$$\hat{y}_i = b_0 + b_1 D_i + b_2 S_i + b_3 (D_i \times S_i)$$

where D = 0 or 1 for the 20g and 40g group, respectively, and S = 0 or 1 for the milk and milk-free LNS group, respectively. Participants who received 10g / day milk-LNS and participants who received no dietary intervention will be excluded in this model. Coefficient b_2 will indicate whether the 20g milk-LNS group differs from 20g milk-free LNS group and linear combination of coefficients b_2+b_3 (Stata's *lincom* command) will tell us whether the 40g milk-LNS group differs from 40g no-milk-LNS group. If $P<0.05$ for coefficient b_2 or b_2+b_3 we will reject the null hypothesis of no differences between groups and use separate groups for 20g and 40g milk and milk-free LNS groups in the subsequent analyses. If $P>0.05$ for coefficient b_2 or b_2+b_3 we will combine milk and milk-free LNS groups in the subsequent analyses.

Another analysis will be done to compare all the LNS groups combined against the control (no dietary intervention) group.

These analyses will primarily be based on principle of *modified* intention-to-treat. The modification is due to six participants who were accidentally allocated to a group other than the one to which they were actually randomized. For each participant, the randomization code was pre-packed and sealed in an individual envelope that was opened and used for group allocation at enrolment. For these 6 individuals, the randomizer made a recording error, i.e. s/he noted down in a data collection form an incorrect group code or wrote the code with unclear handwriting. The incorrect code was later transcribed into the computer software that was used to plan participant visits and allocate interventions. These six participants were told to belong to the erroneously recorded intervention group and they received that intervention throughout the trial – hence they will also be analyzed in that group (rather than the one written on the randomization slip). All randomized participants are eligible to be included in the analyses; only subjects with missing data on an outcome variable will be excluded for the analysis of that outcome.

Besides the intention to treat, per protocol analysis will be done by restricting the analysis to those consuming the supplements for at least 70% of the intervention period.

7 Time points for analyses

All the above analyses will cover the period from enrolment (when the participants are 6 months old) to the end of the intervention (when the participants are 18 months old).

8 Calculation of end-points, presentation of study findings and comparison between groups

8.1 Participant flow

All screened and enrolled infants and the success of the follow up will be described in a flow chart according to CONSORT guidelines (Figure 1). For additional information the drop-out rate between groups will be compared with Fisher's exact test but will not be included in Figure 1.

8.2 Baseline characteristics

Participant characteristics at enrolment will be tabulated by treatment arms (Table 1). Hypothesis testing will be performed for baseline characteristics to give additional information. Methods used for hypothesis testing are also indicated in Table 1, but p-values will not be presented in the final manuscript. Baseline characteristics of refusals and drop-outs will be compared to those who completed the study. P-values for these tests will be shown in the text.

8.3 Risk of serious adverse events

Risk of any SAE, hospitalization and death will be calculated as number of participants having any SAE or hospitalization or death / the total number of participants in each group (Table 2). These events are counted once for each participant, as opposed to total SAEs which was a repeated event in some participants. Risk ratio (95% CI) will be calculated for each intervention group using the control group as reference. Cumulative incidence curve for deaths for each group will be shown in Figure 2.

8.4 Incidence of non-scheduled visits

The total number of non-scheduled visits will be tabulated by intervention group. The incidence of non-scheduled visits will be calculated as the number of non-scheduled visits / the total days of follow up in each group (Table 3). Incidence rate ratio (95% CI) will be calculated for each intervention group using the control group as reference.

8.5 Longitudinal prevalences of self-reported common childhood morbidity symptoms

Longitudinal prevalence is defined as the proportion of days with the symptom in the study period. This will be calculated as total days of each reported morbidity symptom / total number

of days of follow up for each participant in each group to obtain group means. Geometric means ratio (95% CI) will be used to estimate and compare the mean longitudinal prevalences among the intervention and control groups using the control group as the reference. Results will presented by intervention group (Table 4).

8.6 Incidences of common childhood infectious disease episodes

Incidence of episodes for each disease (number of episodes per child / total days of follow up per child) will be presented by intervention group. Incidence rate ratios (95% CI) will be calculated for each intervention group using the control group as the reference. Results will presented by intervention group (Table 5).

9 Hypothesis testing and statistical modelling

Risk of serious adverse events will be analysed using generalised linear modelling (log-binomial family). The Cox regression will be used in the analysis of mortality outcome. Differences in longitudinal prevalences of self-reported symptoms will be analysed and tested using least squares regression. Incidences of non-scheduled visits and morbidity episodes will be analysed using negative binomial regression.

The following is the list of the end-points and the types of modelling to be used for each end-point:

End-point		Statistical modelling
i.	Risk of SAE	Log-binomial regression
ii.	Longitudinal prevalence	Least squares regression
iii.	Incidence of non-scheduled visits	Negative binomial regression
iv.	Incidence of disease episodes	Negative binomial regression

A set non-inferiority margin of 20% will used to test the hypotheses of no marked increase in the risk of SAEs, incidences of non-scheduled visits and morbidity episodes and prevalences of self-reported morbidity symptoms due to the intervention. An increased risk of >20% (risk ratio/incidence rate ratio >1.20) suggests a harmful effect associated with the intervention. The non-inferiority margin of 20% or an incidence difference of 20% has been used in previous

studies on nutritional supplementation and morbidity (Lemaire et al, 2011; Sazawal et al, 2006; Soofi et al, 2013).

The findings of the analysis will lead to one of the following four possible conclusions:

i. If the upper bound of the 95% CI for the risk ratio/incidence rate ratio is <1.20, a conclusion of non-inferiority will be made.
ii. If the lower bound of the 95% CI for the risk ratio/incidence rate ratio is >1.20, a conclusion of inferiority will be made.
iii. If the upper bound of the 95% CI for the risk ratio/incidence rate ratio is <1.00, superiority will be confirmed (suggesting a beneficial effect).
iv. If the upper bound of the 95% CI for the risk ratio/incidence rate ratio is >1.20 and lower bound is <1.20 we will interpret that the study has shown neither superiority nor inferiority. The findings will be considered inconclusive for target group inference.

The results will be plotted on a graph (Figure 3).

10 General notes on statistical methods

10.1 Software

All analyses will be done using Stata version 12. Some analyses and graphs will be done in R-project and Microsoft Office Excel.

10.2 Preparing morbidity data for analysis

Using the weekly morbidity form (Form 18, table) morbidity variables were split to create a new dataset for each morbidity symptom. I created new variables for total days with symptoms and total follow up days in each dataset which were used to calculate prevalence of each morbidity symptom per child. Mean longitudinal prevalences for each group were derived from the individual prevalences. I created morbidity episodes from a combination of symptoms by merging the datasets of self-reported symptoms. Incidences in a group were calculated as the sum of events/episodes across individuals divided by sum of follow-up times across individuals. Preparation of these data was done using STATA and detailed steps taken in preparation of these data were recorded in a STATA do file

10.3 Confidence intervals

The confidence intervals at 95% level (95% CI) will be provided for all the outcomes. I will use log-binomial regression to derive the confidence intervals for SAE risk ratio, geometric means ratio to derive the confidence intervals for mean longitudinal prevalence and negative binomial regression to derive the confidence intervals for incidence of nonscheduled visit and self-reported disease episodes.

10.4 Interaction and effect modification

The tests for interaction will include all variables that could logically modify the effect of nutrition intervention on morbidity. The following variables will be used to test the interactions:

a. Child length for age z-score, weight for age z-score and weight for length z-score at baseline
b. Child sex
c. Maternal education and marital status
d. Proxy for hygiene/water quality
e. Iron status at baseline
f. Anaemia status at baseline
g. Household food insecurity access (HFIA)
h. Seasonality

10.5 Covariate adjustment

The main analysis will be completed and shown in figures and tables without any covariate adjustment.

As a secondary analysis, generalised linear modelling (log-binomial family) will be used to model the possible predictors of SAEs. Least squares regression will be used to model the percentage of days with self-reported morbidity symptoms. Negative binomial regression will be used to model the predictors of non-scheduled visits and disease episodes. Fixed effects will be used to model for age (grouped in 2 age bands) and seasonal variation. All variables which show a statistically significant association with any of the outcomes (at $p<0.05$) will be included in the regression models. The covariates to be included in the models will be derived from the list below:

a. Child length for age z-score, weight for age z-score and weight for length z-score at baseline
b. Child sex
c. Maternal education and marital status
d. Proxy for hygiene/water quality
e. Iron status at baseline
f. Anaemia status at baseline
g. Household food insecurity access (HFIA)
h. Seasonality
i. Age (divided into two bands: <12 months and >12 months)

11 References

Dewey KG, Yang Z, Boy E. Systematic review and meta-analysis of home fortification of complementary foods. *Maternal Child Nutrition* 2009;**5**:283-321.

Kathryn G. Dewey and Lacey M. Baldiviez. Safety of Universal Provision of Iron through Home Fortification of Complementary Foods in Malaria-Endemic Areas. *Adv. Nutr* 2012;**3**:555–559.

Lemaire M, Islam QS, Shen H, Khan MA, Parveen M, Abedin F, Haseen F, Hyder Z, Cook RJ, Zlotkin SH. Iron-containing micronutrient powder provided to children with moderate-to-severe malnutrition increases haemoglobin concentrations but not the risk of infectious morbidity: a randomized, double-blind, placebo-controlled, noninferiority safety trial. *Am J Clin Nutr* 2011;**94**:585–93.

Sazawal S, Black RE, Ramsan M, Chwaya HM, Stoltzfus RJ, Dutta A, Dhingra U, Kabole I, Deb S, Othman MK, Kabole FM. Effect of routine prophylactic supplementation with iron and folic acid on admission to hospital and mortality in preschool children in a high malaria transmission setting: a community-based, randomized, placebo-controlled trial. *Lancet* 2006;**367**:133–143.

Sharieff W., Bhutta Z., Schauer C., Tomlinson G. & Zlotkin S. Micronutrients (including zinc) reduce diarrhoea in children: the Pakistan Sprinkles. Diarrhoea Study. *Archives of Disease in Childhood* 2006;**91**:573–579.

Soofi S, Cousens S, Iqbal S, Akhund T, Khan J, Ahmed I, Zaidi A, Bhutta Z. Effect of provision of daily zinc and iron with several micronutrients on growth and morbidity among young children in Pakistan: a cluster-randomised trial. *Lancet 2013*. Published online.

US Food and Drug Administration. Reporting Serious Problems to FDA. http://www.fda.gov/safety/medwatch/howtoreport/ucm053087.htm. (Accessed on 10th April 2014).

12 Legends to figures

Figure 1: Participant flow in CONSORT recommended format

Figure 2: Cumulative incidence curve for mortality

Figure 3: Graph of non-inferiority, comparing the risk ratios/incidence rate ratios of SAEs, non-scheduled events and care-giver reported morbidity between intervention and control group.

13 Tables

TABLE 1:

Baseline characteristics

Variable	**Control**	**10g**	**20g**	**40g**	**Test**
Number of participants	xxx	xxx	xxx	xxx	-
Mean (SD) age, months	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	ANOVA
Proportion of boys	xx.x%	xx.x%	xx.x%	xx.x%	χ^2
Mean (SD) HAZ-score	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	ANOVA
Mean (SD) WAZ-score	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	ANOVA
Mean (SD) WHZ-score	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	xx.x (x.x)	ANOVA
Mean (SD) haemoglobin, g/L	xxx (x.x)	xxx (x.x)	xxx (x.x)	xxx (x.x)	ANOVA
Proportion with anemia	xx.x%	xx.x%	xx.x%	xx.x%	χ^2
Proportion with iron deficiency	xx.x%	xx.x%	xx.x%	xx.x%	χ^2
Proportion with positive malaria RDT	xx.x%	xx.x%	xx.x%	xx.x%	χ^2

TABLE 2

Risk of SAEs by intervention group

	Results by study group				Comparison between the groups Risk ratio (95% CI)		
Variable	Control	10g LNS	20g LNS	40g LNS	10g LNS vs control	20g LNS vs control	40g LNS vs control
Number of participants	xxx	xxx	xxx	xxx	-	-	-
Number of SAEs	xxx	xxx	xxx	xxx	-	-	-
Number (%) of children who experienced any SAE	xxx (%)	xxx (%)	xxx (%)	xxx (%)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Number (%) of children who experienced a hospitalization	xxx (%)	xxx (%)	xxx (%)	xxx (%)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Number (%) of children who died	xxx (%)	xxx (%)	xxx (%)	xxx (%)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)

TABLE 3

Incidence of non-scheduled visits and clinical diagnoses, by intervention group

	Incidence				Comparison between the groups Incidence rate ratio (95%CI)		
Variable	Control	10g LNS	20g LNS	40g LNS	10g LNS vs control	20g LNS vs control	40g LNS vs control
Number of participants	xxx	xxx	xxx	xxx	-	-	-
Total follow up days	xxxxx	xxxxx	xxxxx	xxxxx	-	-	-
Non-scheduled visits	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Gastroenteritis	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Acute respiratory infections	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Clinical malaria	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Others	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)

Incidence = number of events / total follow up days

TABLE 4:

Longitudinal prevalence of self-reported common childhood morbidity symptoms

	Prevalence				Comparison between the groups Geometric means ratio (95%CI)		
Variable	Control	10g milk LNS	20g LNS	40g LNS	10g LNS v control	20g LNS v control	40g LNS vs control
Total follow up days	xxxxx	xxxxx	xxxxx	xxxxx	-	-	-
Any illness	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Fever	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Cough	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Vomiting	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Rapid breathing	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Difficult breathing	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Nasal discharge	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Diarrhoea	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Bloody stools	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Mucoid stools	xx.x%	xx.x%	xx.x%	xx.x%	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)

Longitudinal prevalence = number of days with symptom / total number of days of follow up

TABLE 5:

Incidence of self-reported disease episodes

	Incidence				Comparison between the groups Incidence rate ratio (95%CI)		
Variable	Control	10g milk LNS	20g LNS	40g LNS	10g LNS v control	20g LNS v control	40g LNS vs control
Total follow up days	xxxxx	xxxxx	xxxxx	xxxxx	-	-	-
Gastroenteritis	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Acute respiratory infections	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Malaria	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)
Others	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xxxx)	x.xx (xx to xx)	x.xx (xx to xx)	x.xx (xx to xx)

Incidence = number of episodes per child / total follow up days