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

Statistical Analysis Plan

Version 02.0 (07.05.2013), includes 3 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**	**Signatures**
Kenneth Maleta, MBBS, PhD	Principal investigator	*Signatures withheld from internet version*
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)
02.0	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)

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) and the incidence of episodes / year of follow-up will be presented by intervention group and SAE categorization, as indicated in Table 4. The negative binomial regression will be used to compare the number of SAE between intervention groups and control groups. Rate ratios (95% CI) for the incidence of any SAE will be calculated for each intervention group (as compared to the control group).

For each participant, SAE data were collected from enrolment (child age six months) until the child had completed the intervention and attended the visit 52 at which point primary outcomes were measured (age 18 months). Some participants were, however lost to follow-up. Time at risk for SAEs for each participant will be calculated by subtracting the date of her enrolment from a date 52 weeks later (the planned date for completing the intervention), the date of withdrawal of consent (active drop-out), or death, whichever of the three occurred first. AE information was

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

For any SAEs and each type of SAE defined in Table 4, the likelihood ratio test will be used to test the global null hypothesis of no differences between groups in incidence of episodes / year of follow-up (Table 4). Global null hypothesis of no differences will be rejected if $P<0.05$. If global null hypothesis of no differences is rejected, the rate ratio and its raw P-values comparing each intervention group versus the control group will be presented in the text.

In addition to calculating the numbers of SAEs per follow-up time and group, there will be another 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) group will be presented as indicated in table 5. 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 the Fisher's exact test.

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 ≤ 0.5 cm
2. circumferences (head, MUAC) ≤ 0.5 cm
3. infant/child weight ≤ 0.1 kg
4. adult weight ≤ 0.1 kg
5. skinfold thickness ≤ 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 “Version history” on page 5 lists the editions made to the different approved versions of the SAP:

15 List of appendixes

Statistical Analysis Plan, Appendix 1: Secondary growth outcomes (added on 17.03.2013).

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).

16 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

17 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.

18 Figures

Figure 1. Participant flow

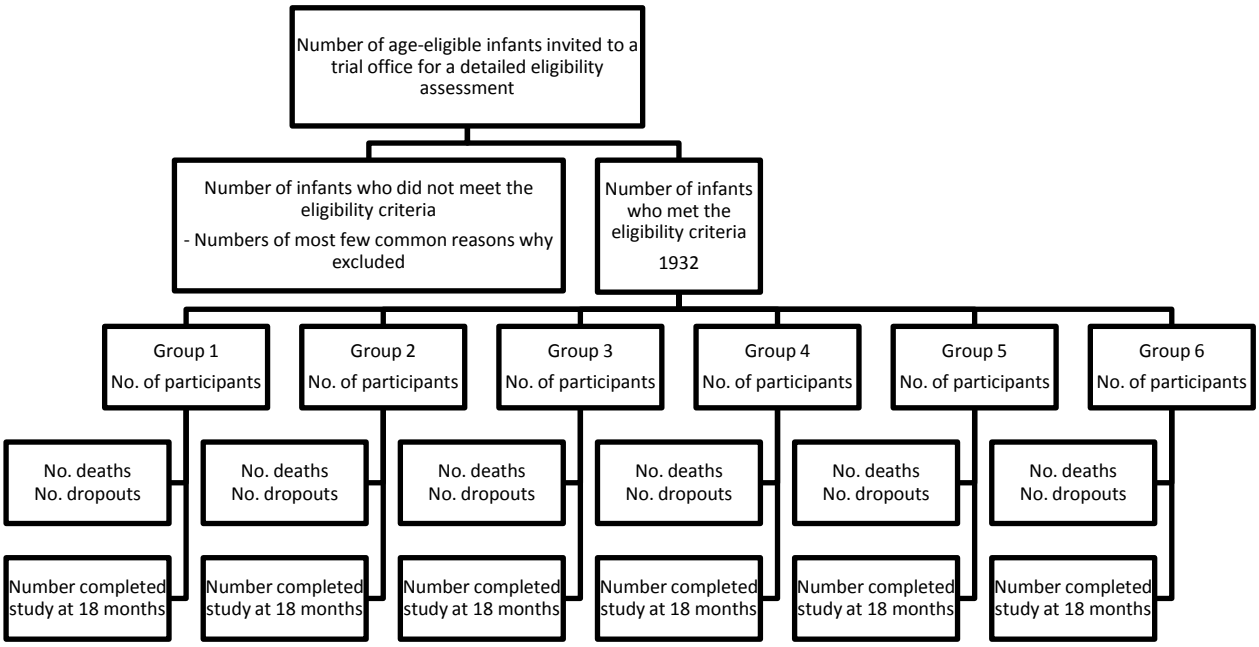


Figure 2. Box-Whisker plots of change in LAZ by groups

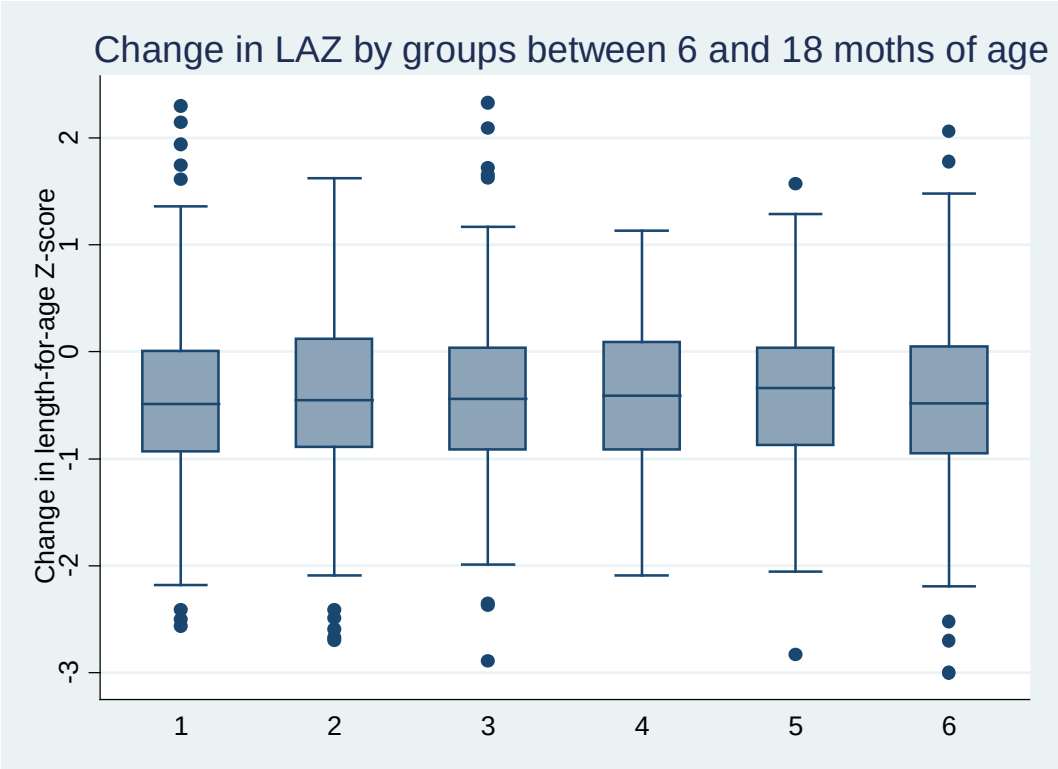


Figure 3. Cumulative incidence curve for mortality, by intervention groups

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 Number of serious adverse events per time at risk per intervention groups

Intervention group	**Control**	**10g milk LNS**	**20g milk LNS**	**20g milk-free LNS**	**40g milk LNS**	**40g milk-free LNS**	**P-value**
Number of months at risk, total (per participant)	xxxx (xx.x)	xxxx (xx.x)	xxxx (xx.x)	xxxx (xx.x)	xxxx (xx.x)	xxxx (xx.x)	
Fatal events, number (incidence per follow-up year)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	x.xxx
Life threatening, number (incidence per follow-up year)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	x.xxx
Inpatient hospitalization, number (incidence per follow-up year)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	x.xxx
Significant disability, number (incidence per follow-up year)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	x.xxx
Other SAE, number (incidence per follow-up year)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	x.xxx
Any SAE, number (incidence per follow-up year)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	x.xxx
Rate ratio (95% CI) between the indicated intervention and the control group	1.00 (Ref)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	xxx (xx to xx)	

Table 5 Proportion of participants with at least one episode of SAE

	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 with at least one episode of SAE / total number of participants in this group (% with SAE)	xx/xx (xx.x%)	xx/xx (xx.x%)	xx/xx (xx.x%)	xx/xx (xx.x%)	xx/xx (xx.x%)	xx/xx (xx.x%)	xx.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 03: Q2.3; 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 03: Q2.3; 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 03: Q2.3; 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 03: Q2.3; 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 03: Q2.3; 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 03: Q2.3; 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

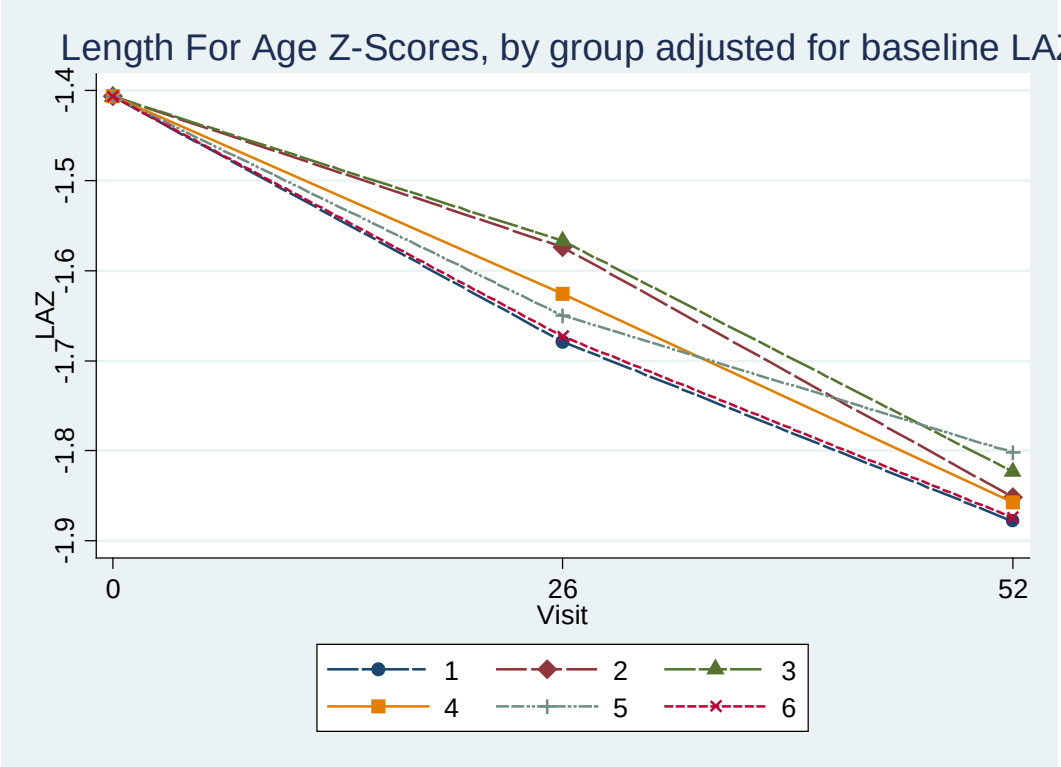
Aickin M, Gensler H. Adjusting for multiple testing when reporting research Results: The Bonferroni vs Holm Methods. *American Journal of Public Health* 1996; 86(5): 726-728.

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

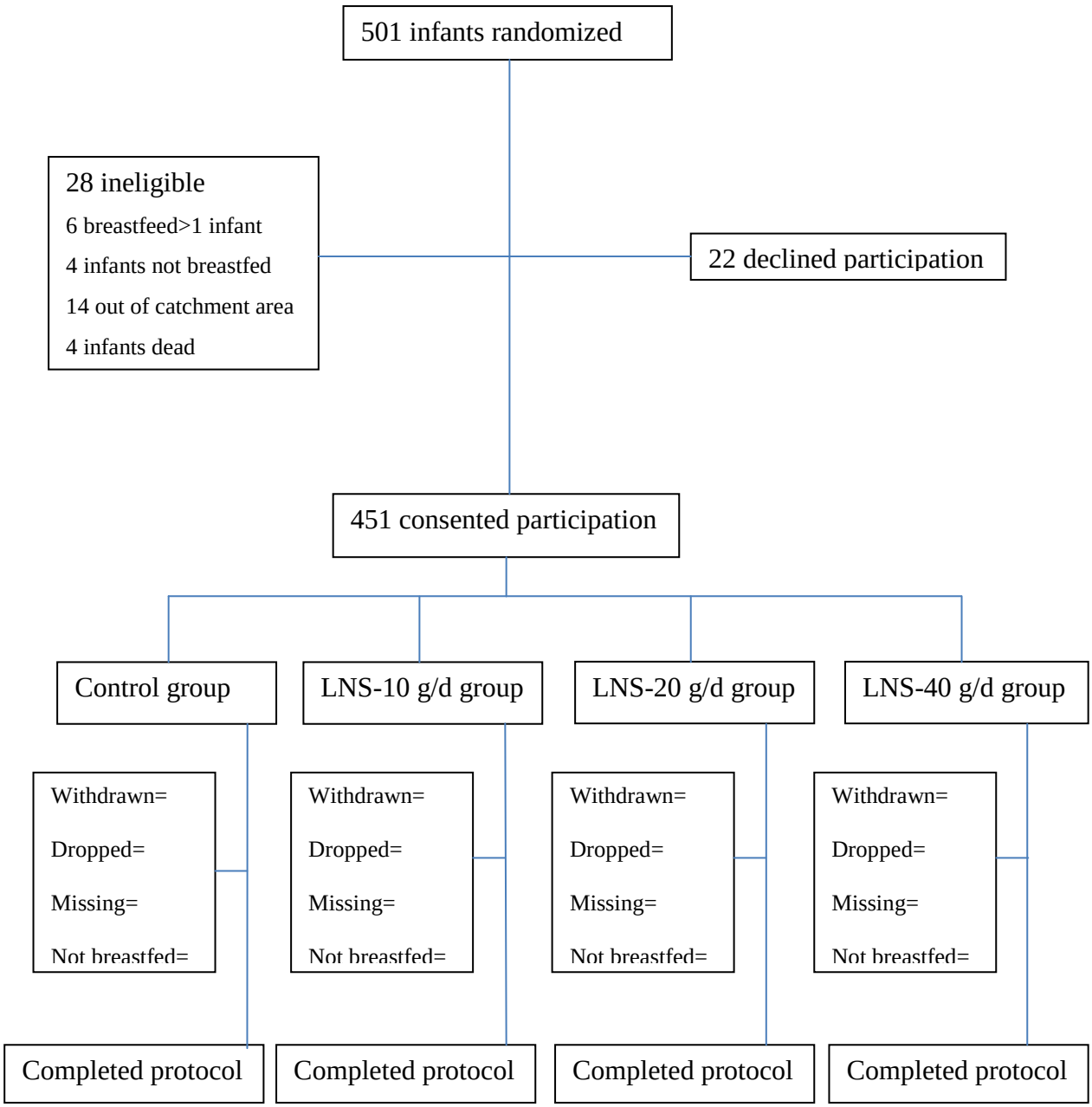
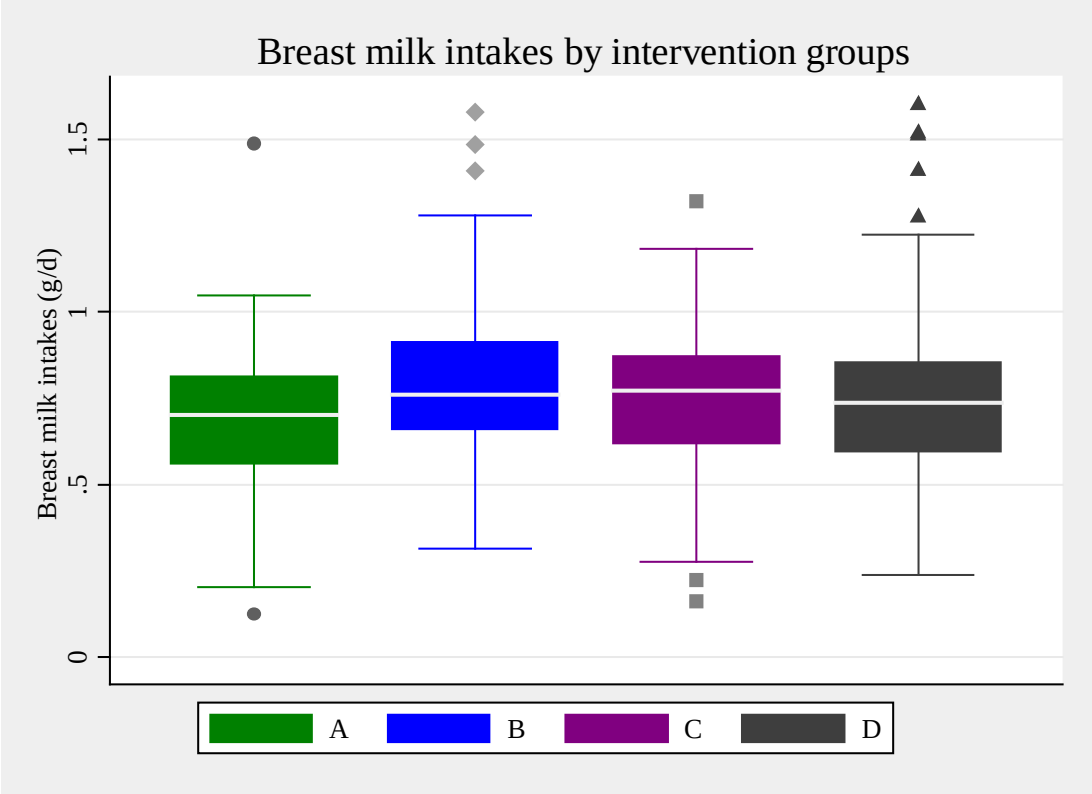


Figure 2. Box-Whisker plots of breast milk intakes by the intervention group



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