Statistical Analysis Plan: iLiNS DYAD-G: The effect of Lipid-based Nutrient Supplement (LNS) supplementation on child morbidity

1. Version History:

Version Number	Version Date	Prepared By	Description of Editions
1.0	November 11, 2016	Rebecca Young	Original Document

2. Overview and study objectives

The primary aim of the study is to compare the effects of three types of nutrient supplements on child growth: 1.) LNS-P&L given to pregnant and lactating women and LNS-20gM provided to children from 6 to 18 months of age, 2.) Multiple Micronutrient tablets (MMN) given to pregnant and lactating women and 3.) Iron and Folic Acid (IFA) tablets given to pregnant women.

3. Description of the study

A more detailed description of the iLiNS-DYAD-G study, including the study population, inclusions and exclusions criteria, blinding, etc. are available in the main statistical analysis plan (iLiNS-DYAD-G Statistical Analysis Plan Version 3, 2014-05-27). Screening, recruitment and enrollment of pregnant women was performed on a rolling basis between December 2009 and December 2011. Screening was performed on women attending prenatal clinics in the Manya Krobo and Yilo Krobo districts in the Eastern Region of Ghana. Eligible and willing women were recruited to participate in the study and randomly selected into one of the three trial arms: 1) daily iron folic acid (IFA) tablets throughout pregnancy and a placebo of low-dose calcium during the first six months of lactation and no supplement for the children born to these women, 2.) multiple micronutrient tables (MMN) taken once a day during both pregnancy and the first six months of lactation and no supplement for the children born to these women or 3.) LNS-P&L during pregnancy and the first six months of lactation; children born to the women in the LNS group also received a child specific formulation of LNS between 6 months of age and 18 months of age.

Morbidity data were collected weekly (± 3 days) at home visits between birth and 18 months of age. Morbidity symptoms were recorded for each day during the previous week.

4. Specific objectives

The specific objective of this study is to compare the following outcomes between the children who received LNS and those who did not receive LNS between the ages of birth to 18 months:

- Incidence of non-scheduled visits made to health facilities due to malaria or other illness
- Longitudinal prevalence of self-reported common childhood morbidity symptoms
- Incidence of common childhood disease episodes

5. Hypothesis to be tested

a.) In children between 6 months and 18 months of age, incidence of non-scheduled visits made to health clinics or hospitals due to illness, including malaria, is not higher in children who received LNS compared to children who received no dietary intervention.
b.) In infants younger than 6 months, incidence of non-scheduled visits made to health clinics or hospitals due to illness, including malaria, is not higher in infants whose mothers received LNS compared to infants whose mothers received MMN or IFA.
c.) In children between 6 and 18 months of age, longitudinal prevalences of self-reported daily morbidity symptoms are not higher in children who received LNS compared to children who received no dietary intervention.
d.) In infants younger than 6 months, longitudinal prevalences of self-reported daily morbidity symptoms are not higher in infants whose mothers received LNS compared to infants whose mothers received either IFA or MMN.
e.) In children between 6 months and 18 months of age, incidence of self-reported episodes of common childhood infectious disease episodes (malaria, gastroenteritis and acute respiratory illnesses) are not higher in children who received LNS than in children who received no dietary interventions.
f.) In infants younger than 6 months, incidence of self-reported episodes of common childhood infectious disease episodes (malaria, gastroenteritis and acute respiratory illnesses) are not higher in infants whose mothers received LNS than in infants whose mothers received MMN or IFA.

6. Definitions of end-points

Non-Scheduled Visits are defined as visits to a clinic, traditional healer, hospital or pharmacy due to morbidity symptoms. Visits are defined as any visit to a clinic, hospital, traditional healer, pharmacy, or private physician regardless of admittance. Non-scheduled visits are a composite of morbidity, access to health care and health seeking behavior. The total number of visits to a health facility made by a participant between the ages of birth to 6 months, birth to 18 months and 6 months to 18 months will be tabulated. Visits made for a well-child appointment will not be included as a non-scheduled visit. *The data will be extracted from Form C1b, questions 23, 24, 25, 26, 30 and 31.*

Self-reported symptoms of any illness for each day were collected at a weekly home visit. These data will be extracted from Form C1b and will include the following symptoms: rashes (Q10), reduced activity due to illness (Q11), poor appetite (Q12), poor stool consistency (semi-liquid, liquid or very liquid, Q13), blood or mucus in stool pathology (Q16), vomiting (Q17), fever (Q18), cough (Q19), rapid breathing (Q20), difficult breathing (Q21) and nasal discharge (Q22).

Diagnoses of malaria, gastroenteritis, and acute respiratory infections are derived from presence of a combination of self-reported symptoms on one or more days. A diarrhea episode is defined as 3 or more loose stools in 24 hours in the presence or absence of other symptoms and with or without blood in the stools. A separate diagnosis of diarrhea will be made if the mother reports that the child has diarrhea, this diagnosis will not be part of the total number of disease episodes definition as it will not be mutually exclusive with diarrhea as reported by the number of loose stools. The

diagnosis of dysentery (a subset of diarrhea cases) is made if there were three or more loose stools per day and there was blood in the stools. The diagnosis of acute respiratory illness (ARI) is made if diarrhea was absent but there was presence of either cough, rapid breathing, difficult breathing, or nasal discharge. Fever was not required for a diagnosis of ARI. Fever episodes in the absence of diarrhea and respiratory symptoms, with or without other symptoms, are categorized as suspected malaria. The rest of the symptoms in the absence of diarrhea, respiratory symptoms or fever are categorized as "others." For all diseases, an episode is defined as the period starting from the day the child had symptoms when preceded by at least 2 days of no symptoms or of no data, and lasting until the last day that the child had symptoms that followed by at least 2 symptom-free days.

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

Incidence of non-scheduled visits

The number of non-scheduled visits to a health facility will be tabulated by intervention group. The incidence of non-scheduled visits will be calculated as the number of non-scheduled visits divided by the total days of follow up in each group (Table 3.). Incidence rate ratio with a 95% confidence interval will be calculated for the LNS and control groups using the control group as a reference.

Longitudinal prevalence 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 divided by the total number of days of follow up for each participant in the LNS versus control groups. Group means will be obtained and the geometric mean ratio with a 90% confidence interval will be used to estimate the mean longitudinal prevalence among the intervention and control groups (Table 4).

Incidences of common childhood infectious disease episodes

Common childhood infectious disease episodes will include Gastroenteritis, Acute respiratory infections (ARIs), Malaria and "other." Incidence of episodes for each disease will be calculated as the number of episodes per child divided by the total days of follow up per child. Incidence of episodes will be presented by intervention group along with the incidence rate ratio and a 90% confidence interval (Table 5).

8. Basis of analysis: Intention to treat and per protocol analysis

- The primary analysis will be by intention-to-treat. That is, results for all children enrolled will be analyzed by the group to which they were assigned regardless of any protocol violations. Data on subjects who were lost to follow-up because of death, travel from the study site, or refusal to continue with the study will be included in the analysis if available.
- A sensitivity analysis will be conducted in which all three treatment groups will be compared for all outcomes. This sensitivity analysis will only consider results unadjusted for covariates. The sensitivity analysis will be performed separately for children in the 0 to 6 month interval and children in the 6 to 18 month interval. If the sensitivity analysis reveals that there are no differences between the IFA and MMN groups, the primary analysis will be a two group comparison between LNS and the combined control group. If the sensitivity analysis reveals that there are pairwise differences between the IFA and MMN groups, then the primary analysis will be a three group comparison.

- At six months, the supplementation in the LNS group switched from a maternal supplement to a child supplement and the mothers in the MMN and IFA groups stopped receiving a supplement. Therefore, the primary analysis will be segmented into two time intervals. The first interval will be between 0 months and 6 months. The second interval will be between 6 months and 18 months. As a secondary analysis, a third interval will be from 0 to 18 months.
- When more than 10% of observations are missing for a dependent variable, we will report the number of observations used in the analysis.
- A per-protocol analysis will be conducted to include only children for whom reported adherence was greater than 50% (in the LNS group). As children whose mothers received MMN and IFA did not receive supplements, adherence data are not available for those two groups. Therefore, predicted adherence will be calculated for children whose mothers were in the MMN and IFA treatment groups and children with a predicted adherence of over 50% will be included in the minimum-adherence per-protocol analysis.
- If MMN and IFA have pairwise differences as determined from the sensitivity analysis, an additional per-protocol analysis will be performed for children whose mothers were not enrolled during the period of the IFA and MMN mislabeling and therefore received the supplement as assigned without any mixed exposure.

9. Hypothesis testing and statistical modelling

As the incidence of non-scheduled visits and the incidence of disease episodes are both count variables, the negative binomial model will be used. It is assumed that the incidence of non-scheduled visits and the incidence of disease episodes both display over-dispersion and zero-inflation. The coefficients derived from negative binomial regression will be interpreted as incidence rate ratios. As some children exited the study before they reached 18 months of age and because some data are missing, children have different numbers of observation days. Therefore, the regression estimated will be adjusted for the number of observation days for each child.

Non-inferiority testing will be used. For the two group comparisons of children supplemented after 6 months, the null hypothesis is that LNS is inferior to the control group which is the combined MMN and IFA arms. For the three group comparison of child outcomes between birth and 6 months, each pairwise comparison will be conducted. In the comparison of LNS to IFA and MMN to IFA, the IFA arm will be considered the control group. In the comparison of LNS to MMN, MMN will be considered the control group.

The non-inferiority margin (δ) will be set as no greater than 20% increase in morbidity in the intervention groups compared to the control group. For each outcome, a 95%, two-sided, confidence interval will be generated for the risk ratio (for outcomes of prevalence of morbidity related symptoms) or incidence rate ratio (for outcomes of incidence of non-scheduled medical visits and incidence of infectious disease). There are four potential conclusions for each confidence interval:

a.) If the upper bound of the 95% CI is greater than δ a conclusion of non-inferiority will be made.
b.) If the lower bound of the 95% CI is less than δ then a conclusion of inferiority will be made.
c.) If the upper bound of the 95% CI is less than 1 (where 1 represents the IRR or RR of the control group) a conclusion of superiority will be made.

d.) If the upper bound of the 95% CI is greater than δ and the lower bound of the 95% CI is less than δ an inconclusive finding will be made.

10. Statistical Methods

Software

All statistical analysis will be performed using SAS version 9.4. R or Microsoft Office Excel will be used for preparation of figures.

Preparation of morbidity data for analysis

Using the daily morbidity data, morbidity variables will be split into separate datasets for each symptom. The number of observation days will be calculated for each symptom, as some children have missing data for some symptoms and not for other symptoms for the same observation day. The individual prevalence for each symptom will be calculated for each child and then mean longitudinal prevalence will be calculated for each group. Morbidity episodes will be calculated as described above. Incidence for each infectious disease episode will be calculated using the definitions described above.

Interaction and effect modification

Tests for significant interactions between the intervention group and baseline covariates will be included for any variable that could logically modify the effect of the intervention on morbidity. Each effect modifier will be considered separately in the regression model to avoid collinearity. Log logistic regression will be performed to model the percentage of days a child has self-reported morbidity symptoms and negative binomial regression will be used to model the number of non-scheduled health visits and disease episodes. The following variables will be tested for potential effect modification:

- Maternal Age
- Child Sex
- Maternal Education
- Baseline Assets Index
- Baseline Housing Index
- Baseline Household food insecurity Index
- Primiparity
- Children in the household under age 5
- Indoor Air Quality

Covariate Adjustment

The main analysis will be completed with no covariate adjustment. These will be labeled the unadjusted models.

Log-linear 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 adjusted treatment effect of non-scheduled visits and disease episodes. All variables which show a statistically significant association with any of the outcomes ($p<0.1$) will be included in an adjusted regression model. The p-value of the treatment group will then reported in these adjusted models. The covariates to be potentially included in the adjusted models are as follows:

- Maternal Height

- Maternal BMI
- Gestational Age at Enrollment
- Maternal Age
- Child Sex
- Maternal Education
- Baseline Assets Index
- Baseline Housing Index
- Baseline Sanitation Index
- Baseline Household food insecurity Index
- Primiparity
- Season at maternal enrolment (Dry season or not)
- Maternal Hemoglobin concentration at baseline
- Children in the household under age 5
- Indoor Air Quality