Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

Study design:
This is a randomized controlled trial involving children ages ~6-9 to 12-15 months from communities in central Guatemala within Wuqu' Kawoq's catchment area (departments of Chimaltenango, Sololá, Sacatepéquez, Suchitepéquez) individually randomized to either the intervention or control group.

- The *control group* will receive the local standard of nutrition care only.
  The standard of nutrition care in Guatemala includes the following clinical care, as determined to be necessary by the MHA medical and nutrition teams: frequent growth monitoring, general nutrition education, parasite treatment, and multiple micronutrient supplementation.
- The *intervention group* will receive an intervention to promote daily egg consumption for a six month period, in addition to the local standard of nutrition care. Specifically, intervention group participants will be provided with enough eggs for the infant to consume one egg, daily, for six months, and also with education on preparation and consumption of eggs.

Inclusion criteria are:
- Infants who are 6.0-9.0 months of age at baseline
- At least one caregiver willing to provide oral informed consent and participate in study activities
- Planned residence in the study area for approximately the next 18 months
- Singleton birth

Exclusion criteria are:
- Infants with moderate to severe acute malnutrition (weight-for-length z-score < -2)
- Infants with severe anemia (Hgb <7 g/dL per WHO guidelines, with adjustments for altitude as necessary)
- Infants with a chronic medical condition that affects growth (e.g., congenital heart disease, genetic condition) as determined by the MHA staff physician
- Infants whose caregivers have cognitive or other impairments that prevent them from being able to provide informed consent or to reliably provide information required for the developmental assessments
- Infants with a known egg allergy
- Infants with recalcitrant, moderate-to-severe atopic dermatitis
- Infants with a history of anaphylaxis or serious allergic reaction to any substance requiring emergency medical care
- Concurrent participation in any other clinical trial

Study sample size:
The study planned to enroll 1200 infants into two study groups: intervention (n=600 children) and control (n=600 children). The study is powered to detect a difference in effect size of 0.15 for growth, assuming a 20% attrition rate over the course of the six-month study. We used a smaller effect size than found in previous relevant literature, because the control group in those studies did not also receive the clinical care that is part of standard of nutrition care in Guatemala, as will be the case in this study.

Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

1200 children were enrolled in the study, and 1088 completed the study (9.3% attrition rate).

Study objectives:

Evaluate the impact of daily egg consumption during the complementary feeding period in addition to the local standard of nutrition care (i.e., intervention group), compared to the local standard of nutrition care alone (i.e., control group) on the following outcomes, in infants that are 6-9 months old at baseline:

1. Child development, as measured by global development scores (*primary outcome*)
2. Growth, as measured by anthropometrics (*secondary outcome*)
3. Diet quality, as measured by the World Health Organization infant and young child feeding indicators (*secondary outcome*)
4. Hemoglobin (*secondary outcome*)

Efficacy outcomes:

The primary endpoint for this study is differences between groups in child *global development* scores, as assessed using the Caregiver Reported Child Development Instruments (CREDI).

Secondary endpoints for this study include changes in:

Growth (WHO 2006 growth standards)

- Weight-for-age Z-score
- Underweight (any [moderate + severe]: Z-score <-2; severe Z-score <-3)
- Length-for-age Z-score
- Stunting (any [moderate + severe]: Z-score <-2; severe Z-score <-3)
- Weight-for-length Z-score
- Wasting (any [moderate + severe]: Z-score <-2; severe Z-score <-3)
- Head circumference-for-age Z-score
- Low head circumference-for-age (any [moderate + severe]: Z-score <-2; severe Z-score <-3)

Diet quality (WHO infant and young child feeding indicators)

- Minimum dietary diversity
- Minimum meal frequency
- Minimum acceptable diet

Hemoglobin (adjusted for altitude)

- Adjusted hemoglobin value (g/dl)
- Anemia (any: <11.0 g/dl; moderate to severe: <10.0 g/dl)

Blinding:

During data collection, field staff collecting data and investigators outside of Guatemala were blinded to group assignment. All investigators involved in data analysis will be blinded to group assignments until

after consensus is reached among the research team on the interpretation and reporting of the results for the primary and secondary efficacy outcomes.

Analysis principles:
The primary analysis will be by intention-to-treat. Results for children enrolled will be analyzed based on initial group assignment regardless of protocol deviations.

For children who are lost to follow-up, all data available will be used as possible. The number of observations used in each analysis will be reported. Mixed effect models with built in tolerance for missing at random data will be used. In addition, pattern mixture models will be used for sensitivity analysis of potentially not missing at random data.

All tests will be two-sided with $\alpha = 0.05$.

Data cleaning and processing:
A field supervisor checked forms for completeness and consistency and queried field agents when it was necessary to make corrections or seek additional information for missing values.

Data were entered directly into the EMR or REDCap. The REDCap database included range checks for some variables.

Data were descriptively analyzed to identify the number of missing values and outliers. Outliers verified as real data with field staff were retained in the dataset; otherwise, outliers were dropped. Skewed variables or variables with retained outliers will be transformed as appropriate. If outliers remain more than 3SD from the mean after transformation, a sensitivity analysis will be done to determine if the outlier(s) are unduly influencing the results. If no suitable transformation can be found to optimize normality and homogeneity of variances, analysis will be done on ranked data.

Z-scores were calculated for the CREDI and growth outcomes using instructions and software/code available from the CREDI developers (https://credi.gse.harvard.edu/credi-materials) and from WHO (https://www.who.int/tools/child-growth-standards/software).

Scores were calculated for family care indicators following practices used by Hamadani et al. (10.3329/jhpn.v28i1.4520). Scores for food security were calculated using instructions and software/code available from FAO (https://www.fao.org/in-action/voices-of-the-hungry/analyse-data/en/). Poverty score was calculated using instructions from Schreiner (2016).

Hemoglobin values were adjusted for altitude.

Missing data

Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

Efficacy outcomes

Note: may also be included as covariates in models for other outcomes

Variable [variable name]	**Source**	**Missing Data – Baseline N=1200**	**Missing Data – Follow-up N = 1200**	**Missing Data – Follow-up N=1088 (completers)**
Overall CREDI score [baseline: CREDI_Z_OVERALL; follow-up: CREDI_Z_OVERALL_f]	Calculate from development form	None	N=112	None
Weight-for-age Z-score [baseline: waz, endline: waz_f]	Calculated from measurements on anthro form	N=2	N=132	N=20
Length-for-age Z-score [baseline: haz, endline: haz_f]	Calculated from measurements on anthro form	None	N=114	N=2
Weight-for-length Z-score [baseline: whz, endline: whz_f]	Calculated from measurements on anthro form	N=2	N=133	N=21
Head circumference-for-age Z-score [baseline: _zhc, endline: _zhc_f]	Calculated from measurements on anthro form	N=1	N=116	N=4
Minimum dietary diversity [baseline: mdd, endline: mdd_f]	Feeding indicators form	N=85	N=243	N=131
Minimum meal frequency [baseline: mmf, endline: mmf_f]	Feeding indicators form	N=39	N=154	N=42
Minimum acceptable diet [baseline: mad, endline: mad_f]	Feeding indicators form	N=85	N=243	N=131
Infant hemoglobin [baseline: hb_bl, endline: hb_f]	Calculated from hemoglobin measurements on anthro form	None	N=128 One outlier (21.5) recoded to missing	N=16

Potential covariates/confounders

Variable [variable name]	**Source**	**Missing Data – Baseline**	**Missing Data – Follow-up**	**Missing Data – Follow-up**

Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

		N=1200	**N = 1200**	**N=1088 (completers)**
Child sex [sex]	Demographic form – q2.2	None	N/A	N/A
Marital status [civil]	Demographic form – q4.1	N=1	N/A	N/A
Maternal education [school_y]	Demographic form – q4.5	None	N/A	N/A
Maternal literacy [alfabet]	Demographic form – q4.4	None	N/A	N/A
Maternal age [mom_age; mom_age2 (use only when caregiver is not mom)]	Demographic form – q4.2	None	N/A	N/A
Household size [n_household]	Demographic form – q5.1	None	N/A	N/A
# of kids in household under age 5 [child5y; n_child5y (use for analysis)]	Demographic form – q4.6 + 4.7	N=14 Value of 13 confirmed as real (adoption)	N/A	N/A
Poverty score [povertyscore]	Demographic form – q5.2-5.10	N=7	N/A	N/A
Water source [water_source]	Demographic form – q6.1	None	N/A	N/A
Water treatment [water_tx]	Demographic form – q6.3	None	N/A	N/A
Food security [raw_fies]	Demographic form – q7.1-7.8	N=7	N/A	N/A
Family care indicators – play materials [baseline: playmaterials_score; endline: playmaterials_score_f]	Family care indicators form	N=1	N=114	N=2
Family care indicators – play activities [baseline: playactivities_score; endline: playactivites_score_f]	Family care indicators form	None	N=113	N=1
Preferred language [language]	Demographic form – q1.12	None	N/A	N/A
Birth order [birth_order]	Demographic form – q3.1	None	N/A	N/A

Prenatal care during child's pregnancy [pren_care]	Demographic form – q3.2	N=1	N/A	N/A
Prematurity (<37 weeks) [preterm_cat]	Demographic form – q3.4 + q3.5	None	N/A	N/A
Child age [baseline: age_bl, endline: age_f]	Demographic form – q2.4	None	N=112	None
Maternal physiological state (pregnant and/or lactating) [est_fis]	Demographic form – q1.11	None	N/A	N/A
Low birthweight (low_birthwt)	Demographic form – q3.8	N=60	N/A	N/A
Minutes to water source [water_time: source within house or not; water_time1: time to water source if not in the home]	Demographic form – q6.2	water_time: N = 5 water_time1: missing for 2 of 544 households that had to travel for water High values (e.g., 6 hours) confirmed as real	N/A	N/A
Assessment staff [baseline: inv_id_bl; endline: inv_id_f]	Demographic form – q1.5	None	N=112	None
Month of outcome assessment (baseline: date_int_month, endline: date_int_month_f)	Demographic form – q1.1	None	N=112	None
Self-reported adherence to egg intervention	EMR	Decided not to use due to poor data quality.		
# of visits during study	EMR	N=56		
Morbidity – total diarrhea episodes (yes/no) reported at home visits	EMR	Decided not to use due to poor data quality.		
Morbidity - total respiratory episodes (yes/no) reported at home visits	EMR	Decided not to use due to poor data quality.		

Morbidity - total fever episodes (yes/no) reported at home visits	EMR	Decided not to use due to poor data quality.
Total chispitas packets consumed by child every week – reported at home visits	EMR	Collected at six visits during the study – missing data are timepoint specific. Chispitas_1: N=58 Chispitas_2: N=91 Chispitas_3: N=112 Chispitas_4: N=126 Chispitas_5: N=165 Chispitas_6: N=419
Total eggs consumed by child child in the last month reported at home visits	EMR	Collected at six visits during the study – missing data are timepoint specific. monthly_eggconsump_1: N=90 monthly_eggconsump_2: N=116 monthly_eggconsump_3: N=128 monthly_eggconsump_4: N=145 monthly_eggconsump_5: N=168 monthly_eggconsump_6: N=418

Statistical analysis

Descriptive analysis

The following baseline characteristics will be descriptively summarized by treatment group (Table 1) and for completers/non-completers (supplemental table):

Maternal/household characteristics:

- Maternal/caregiver age
- Maternal education level
- Maternal literacy
- Maternal marital status
- Maternal physiological state
- Preferred language
- Household size
- Number of children <5 years old
- Poverty score
- Water source
- Minutes to water source
- Water treatment
- Food security status

Child characteristics:

- Sex

Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

- Child age
- Birth order
- Premature (yes/no)
- Low birthweight (yes/no)
- Prenatal care (yes/no)
- Baseline CREDI Z-score (overall and by domain)
- Family care indicators – play materials
- Family care indicators – play activities
- Baseline anthropometrics
    - Weight
    - Length
    - Head circumference
    - Length-for-age Z-score
    - Weight-for-age Z-score
    - Weight-for-length Z-score
    - Head circumference-for-age Z-score
    - Stunting
    - Wasting
    - Underweight
- Baseline dietary quality
    - Minimum dietary diversity
    - Minimum meal frequency
    - Minimum acceptable diet
- Baseline hemoglobin
    - Continuous value
    - Anemia
- Reported consumption of *chispitas* in the previous week at each visit timepoint (n [%] reporting no consumption, median [IQR])
- Reported consumption of eggs in the previous month at each visit timepoint (n [%] reporting no consumption, median [IQR])

The number of missing values will be reflected in the table. Frequencies and percentages will be used to summarize categorical data, with percentages calculated based on the number of participants with available data. Continuous variables will be reported using mean/SD or median/IQR, depending on the skewness of the data.

Statistical tests will not be used to compare the treatment groups based on these baseline variables. However, differences will be descriptively evaluated, and variables for which there is an indication of imbalance between the groups will be included in models.

<u>Bivariate analysis</u>

Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

Relationships between each of the following variables and the other variables will be evaluated to assess the potential for collinearity:

Maternal/household characteristics:
- Maternal/caregiver age
- Maternal education level
- Maternal literacy
- Maternal marital status
- Maternal physiological state
- Preferred language
- Household size
- Number of children <5 years old
- Poverty score
- Water source
- Minutes to water source
- Water treatment
- Food security status

Child characteristics:
- Sex
- Child age
- Birth order
- Premature (yes/no)
- Low birthweight (yes/no)
- Prenatal care (yes/no)

Relationships between each of the following variables and each of the efficacy outcomes will be assessed to identify covariates for potential inclusion in models ($p<0.10$), after considering potential for collinearity:

Maternal/household characteristics:
- Maternal/caregiver age
- Maternal education level
- Maternal literacy
- Maternal marital status
- Maternal physiological state
- Preferred language
- Household size
- Number of children <5 years old
- Poverty score
- Water source

- Minutes to water source
- Water treatment
- Food security status

Child characteristics:
- Number of study visits
- Sex
- Child age
- Birth order
- Premature (yes/no)
- Low birthweight (yes/no)
- Prenatal care (yes/no)
- Family care indicators – play materials (CREDI models)
- Family care indicators – play activities (CREDI models)
- Baseline and endline anthropometrics
    - Length-for-age Z-score (CREDI, diet quality, and hemoglobin models)
    - Weight-for-age Z-score (CREDI, diet quality, and hemoglobin models)
    - Weight-for-length Z-score (CREDI, diet quality, and hemoglobin models)
    - Head circumference-for-age Z-score (CREDI, diet quality, and hemoglobin models)
- Baseline and endline dietary quality
    - Minimum dietary diversity (CREDI, growth and hemoglobin models)
    - Minimum meal frequency (CREDI, growth and hemoglobin models)
    - Minimum acceptable diet (CREDI, growth and hemoglobin models)
- Baseline and endline hemoglobin (CREDI, growth, and diet quality models)
- Reported consumption of *chispitas* in the previous week at each visit timepoint
- Reported consumption of eggs in the previous month at each visit timepoint

The following variables will be considered as potential effect modifiers in each model:
- Baseline value for the outcome being analyzed (continuous variable or categorical considering the median)
- Baseline LAZ < -1.5 for CREDI, diet quality, and hemoglobin models
- Child sex
- Prematurity (yes/no)
- Low birthweight (yes/no)
- Household poverty score (categorical considering the median)
- Household food insecurity (food insecure versus not)
- Maternal education level

<u>Models</u>

Effects of Complementary Feeding of Eggs on Infant Development and Growth in Guatemala: The Saqmolo Study (Saqmolo)

Statistical analysis plan
Version 1.19.2024

The analysis will test the null hypothesis of no difference in the efficacy outcomes between the two study groups. Multilevel generalized linear or logistic regression models will be constructed as appropriate to the efficacy outcome.

When possible, change scores will be calculated and two-level models will be fit with community (variable "village") as the cluster. Otherwise, all models will have three levels for the repeated measures of participants nest in community.

Regression models will include covariates/confounders as appropriate based on bivariate analysis of relationships between variables and consideration of biological plausibility/review of the literature. There may be a different set of covariates for each outcome.

The effects of potential effect modifiers will be assessed with an interaction term. Significant interactions ($p < 0.10$) will be further examined with stratified analyses, estimation of separate regression lines, or estimation of adjusted means at key points of the covariate, to understand the nature of the effect modification.

Estimates of treatment effects will be accompanied by a 95% confidence interval.