Statistical analysis plan (SAP)[1]

Joint Interventions to Improve Birth Outcomes and Nutrition: the JIBON trial

Section 1 – Administrative information

Title: Joint Interventions to Improve Birth Outcomes and Nutrition: the JIBON trial

Trial registration number: NCT05855551 (clinicaltrials.gov)

SAP Version: 1.0; September 25, 2024

SAP revision history: N/A

Roles and responsibilities

- Jef L Leroy, PhD; Shalini Roy, PhD: principal investigators (International Food Policy Research Institute)
- Moira Angel, MSc: investigator (International Food Policy Research Institute)
- Leila L Larson, PhD; Edward A Frongillo, PhD: investigator (University of South Carolina)
- Fahmida Akter, MSc: PhD student (University of South Carolina)

[1] We follow the guidance provided by Gamble et al. (64)

Section 2 – Introduction

2.1 Background and rationale

Maternal undernutrition is a global public health problem with far-reaching effects for both mothers and infants (1). Poor maternal nutrition negatively affects fetal growth and development. Both micro and macro-nutrients are required for the physiological changes and increased metabolic demands during pregnancy, including fetal growth and development. Inadequate intake of vitamins and minerals is known to negatively affect the health, function, and survival of the mother and fetus. Maternal iodine deficiency is associated with delays in neural, intellectual and physical development; folate deficiency is associated with a higher risk of neural tube defects; vitamin A deficiency in mothers can lead to night blindness; iron deficiency anemia is believed to lead to low birthweight and increased perinatal mortality (1).

Women in Bangladesh have poor diets and are struggling to meet their nutrient requirements, especially during pregnancy and lactation when requirements are higher. Women's diets are generally monotonous: rice accounts for as much as 77% of women's energy intake, with oil contributing 7%, vegetables 6%, and fish 3% (2). In a micronutrient survey conducted in 2011/12 in Bangladesh, up to 26% of non-pregnant and non-lactating women were found to be anemic, and a large share suffer from deficiencies in micronutrients such as zinc (57%), iodine (42%), iron (7%), vitamin B12 (6%), and vitamin A (5%) (3). Results from a nationally representative survey of rural Bangladesh conducted by IFPRI in 2011/12 suggested inadequate intake of several nutrients—iron (65% of required intake), zinc (22%), and vitamin A (97%)[2]—among women of reproductive age, as well as low intakes of energy (47% of women had energy intakes below estimated energy requirements [EER](this draft paper uses data from IFPRI's 2011/12 Bangladesh Integrated Household Survey)(4). Among *pregnant* women, the National Low Birth Weight Survey (NLBS) 2015 reported a prevalence of anemia of 55% (5). A recent dietary survey of 600 pregnant women from four of Bangladesh's poorest districts also documented low energy intakes (46% < EER and 28% < 85% EER) and alarmingly high probabilities of inadequate micronutrient intakes (94% for iron, 75% for zinc, 77% for vitamin A, 93% for calcium, 87% for folate, and 66% for vitamin B12) (6). The excessively high probabilities of inadequate nutrient intakes found in this sample of pregnant women (compared to those in all women of reproductive age from the national survey) may be due in part to the higher nutrient requirements—and inability to meet them—during pregnancy and to the low socioeconomic status of women in this sample. The large disparities that exist between socioeconomic groups, education levels, and urban and rural areas in Bangladesh are well documented and affect many nutrition and diet indicators and their determinants (3,7,8).

Maternal undernutrition during pregnancy is associated with a range of adverse birth outcomes, including stillbirths, preterm births, low birthweight, and small-for-gestational-age (SGA) neonates (1), all of which remain unacceptably high in Bangladesh (9–11). Estimates from 2010 show that 40% of children born alive in Bangladesh were small for gestational age, 22% had low birthweight, and 14% were premature (12). According to the 2015 NLBS survey, low birthweight affected roughly 22.6% of newborns (5). Maternal nutrition in Bangladesh deserves urgent attention in order to protect the survival, health, and well-being of mothers and their young children.

The Bangladesh National Social Security Strategy (NSSS) called for the development of a national Mother and Child Benefit Program (MCBP) to improve the nutrition and health of women during pregnancy and their children up to age 4. The Ministry of Women and Children Affairs (MOWCA) combined two existing

[2] The discrepancy between reported iron deficiency from biomarkers and probability of inadequate intake from dietary surveys may be explained by the well-documented high levels of iron present in groundwater (tube well) in the country (3,65). For vitamin A, the discrepancy is likely to be related to seasonal differences, with plant-based vitamin A being present in large concentrations in highly seasonal fruits. Additionally, both dietary surveys were conducted in rural areas, whereas the micronutrient survey (3) was conducted in rural, urban, and urban slum areas.

cash-based safety nets–the rural Maternity Allowance (MA) and urban Lactating Mother Allowance (LMA) programs–to create the MCBP.

The MA program provided 500 BDT (USD 5.95) per month to poor pregnant women to improve access to nutritious food. Benefits started at enrollment (as early as the first trimester of pregnancy) and were provided for 24 months; women were enrolled once a year in July. Starting in the 2018/19 program cycle, key program improvements included better targeting (using widespread promotion of the program and clear and transparent eligibility criteria), monthly enrollment of eligible pregnant women, timely payments every month for 36 months, and a larger cash transfer (800 BDT per month). To help in improving the design of the MCBP, this study tests different packages of interventions (consisting of different combinations of cash transfers, food transfers, and behavior change communication) using a cluster-randomized study design.

Social protection provides a promising platform on which to leverage improvements in nutrition at scale (13). Current evidence specifically on the impacts of social protection on birth outcomes, however, is limited: few studies have been conducted and some of these studies suffer from methodological limitations (14). The study will contribute to filling this knowledge gap. An additional motivation for the study is provided by the recent WHO 2016 Antenatal Care Guidelines (15). The guidelines call for studies on the effectiveness of alternatives to providing energy and protein supplements to pregnant women (which is recommended in undernourished populations). Studying the effectiveness of providing combinations of food and cash will help build this evidence base. A third reason to conduct the study is that both food transfers and cash transfers are commonly used policy instruments in Bangladesh, and the choice of intervention components to scale up in the MCBP will be guided by the findings from this pilot study. The study findings will thus be highly policy relevant.

2.2. Objectives

The study has two specific objectives:

- Assess the impact of adding a food basket targeted at pregnant women to the standard MCBP program package (which provides cash and behavior change communication (BCC)) on a set of outcomes (see Section 6)
- Assess the impact of adding additional cash targeted at pregnant women to the standard MCBP program package (which provides cash and BCC) on a set of outcomes (see Section 6)

Section 3 – Study methods

3.1. Trial design

A three-arm cluster-randomized, non-blinded, community-based, mixed cross-sectional and longitudinal trial is used to estimate the effectiveness of the intervention. The unit of randomization is a grouping of 3 wards[3]. Each ward group is randomly assigned to one of three study groups, all of which receive the improved maternity allowance program:

- **Arm 1 – "standard" package, i.e., base cash and nutrition behavior change communication (BCC)**: pregnant women receive 800 BDT each month and group-based BCC on nutrition with a focus on how to improve their dietary intake during pregnancy.
- **Arm 2 – "standard" package + food basket**: pregnant women in this arm receive the base cash of 800 BDT each month, group-based BCC, and in addition a monthly food basket composed of rice, lentils, and oil valued at approximately 1,000 BDT.
- **Arm 3 – "standard" package + top-up cash**: pregnant women in this arm receive the base cash of 800 BDT each month, group-based BCC, and in addition receive a monthly additional cash transfer of 1,000 BDT instead of the food basket.

The study is a hybrid of a cross-sectional and longitudinal design. It follows women from their second trimester of pregnancy to around 2 months after childbirth. A census is conducted to identify households with women who met the program's eligibility criteria (see Section 5.2). The study then uses a total of 6 three-month long data collection waves. Women in the 2nd trimester of pregnancy (based on self-reported last menstrual period[4] and ultrasound reports and antenatal care cards when available) were enrolled in the study in the first four waves (**Figure 1**). Household and individual data are collected at enrollment. In the wave following enrollment, data are collected on the same women who are now in the 3rd trimester. Third trimester data collection is conducted at around 34 weeks of gestation (± 1 week). Subsequent waves are used to assess newborn weight and length (collected within 72 hours after child birth) and to assess anthropometric and other outcomes at around 2 months postpartum (61 days ± 1 week). In the final follow-up round (1 month in length), women in the 3rd trimester of pregnancy not captured in wave 5 are surveyed. In addition, the newborn survey is conducted. The number of survey rounds is based on the sample size calculations presented below and field reports regarding the actual number of study subjects enrolled and successfully followed up.

[3] The administrative structure in Bangladesh consists of divisions (8), districts (64), subdistricts or upazilas (491), and unions. Up to about 15 years ago, each union consisted of 3 wards. Since then, the ward system was changed to 9 wards per union, each identified by a number (i.e., from 1 to 9). The study uses the old definition of wards (i.e., 3 per union) as the unit of randomization.

[4] Pregnancy trimesters are defined as follows: 1st trimester: 2 to 13 weeks of pregnancy (meaning, weeks since first day of the last menstrual period); 2nd trimester: 14 to 26 weeks of pregnancy; 3rd trimester: 27 weeks of pregnancy until delivery.

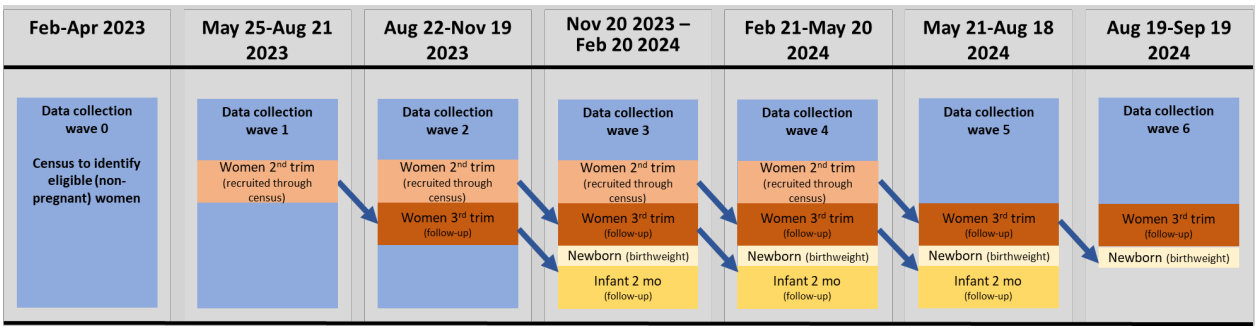

Figure 1: Study design (note: dates are approximate)

3.2. Randomization

The six study upazilas (**Table 1**) are located in distinct parts of the country and are quite heterogeneous. Stratifying clusters prior to randomization is used to ensure a more balanced distribution of cluster-level covariates between study arms. Each upazila is a stratum. In each stratum, the largest number of ward groups that is a multiple of 3 is randomly selected, resulting in a total of 144 ward groups. Each ward group is then be assigned a random number from a uniform [0,1] distribution using Stata. Stratum-specific tertiles are created using this random number and clusters are assigned to the treatment arm that corresponds to their tertile number. Due to the nature of the intervention, trial assignment is not masked.

Table 1: Study upazilas

Division	District	Upazila
Sylhet	Sylhet	Gowainghat
Khulna	Bagherat	Chitalmari
Rajshahi	Rajshahi	Godagari
Rangpur	Rangpur	Gangachara
Barisal	Barisal	Mehendiganj
Mymensingh	Netrakona	Mohanjganj

3.3. Sample size

Sample size requirements were calculated for the primary study outcomes using the Hayes and Bennet's formulas (16,17), using the following parameters: a 0.05 probability of a type I error, power of 0.90, and 48 clusters per treatment arm. Note that that the power of 0.90 (higher than what is typically used) is used to minimize the risk of not detecting a program impact by underpowering the study. Adjustments were further made for intra-cluster correlation (since the study is cluster randomized). The outcome-specific intra-cluster correlation (ICC) coefficients are provided below. The estimated size of the eligible population in the 6 study upazilas was used to calculate the coefficient of variation of cluster sizes (equal to 0.424). We further assumed that baseline covariates provide a small amount of explanatory power of the study outcomes (correlation of 0.1).

We did not adjust the sample size for multiplicity (i.e., carrying out many comparisons). Our thinking is provided in Section 4.1.

We first calculated the required sample size for gestational weight gain and then calculated what minimal detectable effects can be observed with this sample for the other primary outcomes.
The literature provides very limited guidance on the expected effect of the planned intervention packages on **gestational weight gain**. The interventions, through their intensive BCC and provision of cash or food basket, are most akin to the provision of supplements to increase energy and protein intake during pregnancy. In Ghana, lipid-based nutrient supplement (LNS) provided during pregnancy increased gestational weight gain in the second and third trimester by 10 g per week (p=0.09) when compared with the provision of iron-folic acid (IFA) or multiple micronutrient supplements (MMN)(22). In Bangladesh, the effect of LNS (as compared to IFA) on pregnancy weight gain was limited to multiparous women over 25 years of age: an effect of 34 g per week was found (P=0.001)(23). The Women's First Trial which conducted in 4 countries (Democratic Republic of the Congo, Guatemala, India, and Pakistan) found that LNS (compared to control) increased gestational weight gain by 22 g per week (own calculations based on reported values; no p values)(24). Finally, a systematic review and meta-analysis of the effect of balanced energy and protein supplementation on gestational weight gain estimated an average effect of 19 g per week (95% CI: -1.81,39.07)(25). The quality of the evidence, however, was assessed as being of very low quality.

Given the limited evidence from the literature, we based the minimum detectable effect on the difference between observed weight gain and weight gain recommended by the Institute of Medicine (26). Recent studies in Bangladesh found that poor women gain only around 60% of the currently recommended weight gain (adjusted for BMI status)(27,28). We powered the study to detect an effect of 25 g per week, which is equivalent to a 9 to 10% increase relative to current weight gain. Using a standard deviation of weekly gestational weight gain of 161g (average of 4 published studies conducted in Bangladesh (27–30)) and an ICC of 0.01, this requires a sample size of 1,152 women per study arm. The ICC of 0.01 corresponds to a design effect of 1.35. We did not find information on the ICC of gestational weight gain in the literature. Using birth outcome studies, we find that Christian et al. used a design effect of 1.2 in Nepal (31) and West et al. a design effect of 1.15 in a study in Bangladesh (32). An ICC of 1.35 thus appears to be a conservative estimate.

A number of adjustments to these calculated numbers were made. We expect 8.3% fetal loss and 2.3% loss to follow-up between the 2nd and 3rd trimester (33). In order to get the required 1,152 pregnant women per study arm (3,456 women total), we need to enroll 1,286 pregnant women in each arm (3,858 total). For the outcomes measured in newborn children, we assumed 1.7% infant death and 5% loss to follow-up (33). We thus assumed a total of 1,076[5] newborns in each arm (3,227 total). Any additional outcome-specific adjustments to sample sizes are mentioned below.

The detectable differences for the other primary outcomes <u>measured during gestation</u> are as follows:

- **Dietary intake measures** – Using an ICC of 0.05, the sample size calculated for gestational weight gain allows us to observe an effect size of 0.21. We used previously collected data on diet during pregnancy to estimate the ICC (34). The ICCs for energy and protein intake were 0.02

[5] Due to the need to use integral numbers for the number of observations per cluster, we used 22 observations per cluster in the sample size calculations or 22 x 48 = 1,056 observations per arm. Enrollment is targeted to reach the 1,076 per arm which means that the actual observable differences are slightly smaller than what is shown here.

and 0.03 respectively. The ICC for the probability of adequacy of individual micronutrients ranged from 0.00 to 0.27 but was imprecisely estimated for many micronutrients. This might be due to the limited number of clusters (n=18) in the study. The ICC of 0.05 used here is equivalent to a design effect of 2.37.

Using the SD from a recent study in pregnant women in Bangladesh (SD of 831 kcal for energy and 26.8 g for protein) (34), our study is powered to detect effects of 174 kcal for energy and 5.6 g for protein. To assess whether the energy and protein impacts are reasonable in size, we compare them to the size of the food basket that women in one of the study arms receive, to the evidence from the literature, and to the energy and protein requirements during pregnancy. Assuming that the food is distributed equally among all family members (**Table 1**), we have a potential impact of the food basket on energy intake of 354 kcal per person per day and on protein intake of 11 g per person per day, i.e., approximately double the minimal detectable difference calculated above. A recent systematic review of interventions to increase energy and protein intake during pregnancy identified only 3 studies that looked at dietary intake as an outcome (25). These 3 studies assessed the impact of nutrition education vs no nutrition education (i.e., no supplements or food transfers were provided). The estimated effect from the meta-analysis was 105.6 kcal for energy and 7.0 g for protein. This is smaller than the minimal detectable effect calculated above, but the reviewed interventions did not provide cash or food transfers. The expected impact of the Jibon intervention on dietary intake is thus larger. Finally, the additional energy requirements during pregnancy are 340-350 kcal/d in the second and 452-500 kcal/d in the third trimester (35,36). Additional protein requirements for the second half of pregnancy are estimated at 0.30 g/kg body weight/d or 15.3 g /d for an average body weight of 51 kg (34,37). The minimal detectable effect estimated above is considerably smaller than the additional energy and protein requirement for pregnancy. In sum, the effects we are powered to detect (174 kcal for energy and 5.6 g for protein) are well within the range of the food basket's possible contribution to the diet, evidence from the literature, and the energy and protein requirements during pregnancy.

A recent study in pregnant women in Bangladesh showed that the mean probability of adequacy (MPA) was 0.41 ± 0.11 (34). We are thus able to detect an impact of 2.3 percentage points in MPA or a 5% change.

COVID-19 Scenario

Fieldwork might be disrupted due to a COVID-19 outbreak. As a consequence, not all survey rounds may be completed as planned. We thus calculated the minimal detectable effect for the dietary outcomes if only one- or two-thirds of the planned interviews can be conducted. Dietary date on one-third of the pregnant women (i.e., 384 women per arm) allows us to observe an effect size of 0.28, and effects of 233 kcal for energy, 7.5 g for protein, and 3.1 percentage points for the mean probability of adequacy. With two-thirds of the sample, these values change to 190 kcal for energy, 6.13 g for protein, and 2.5 percentage points for the mean probability of adequacy.

- **Length of gestation**: Using an SD of 18.1 days (38), an ICC of 0.01 and the other parameters as defined above, we find that the study is powered to detect an impact on the length of gestation of 2.77 days.

- **Preterm birth**: Assuming a prevalence of preterm birth of 20% in the comparison arm (38) and using an ICC of 0.01, the study's sample size is powered to detect a difference of 6 percentage points in preterm birth.
- **Prevalence of emotional violence towards the pregnant woman:** Assuming a prevalence of physical violence 40.0% (prevalence estimate from the Transfer Modality Research Initiative project or TMRI) and using an ICC of 0.01, the study's sample size provides the statistical power necessary to detect a reduction of 7 percentage points in the occurrence of physical violence.
- **Prevalence of controlling behaviors towards the pregnant woman:** Assuming a prevalence of physical violence 79.5% (prevalence estimate from the Transfer Modality Research Initiative project or TMRI) and using an ICC of 0.01, the study's sample size provides the statistical power necessary to detect a reduction of 7 percentage points in the occurrence of physical violence.
- **Prevalence of physical violence towards the pregnant woman**: Assuming a prevalence of physical violence of 19.9% (prevalence estimate from the Transfer Modality Research Initiative project or TMRI) and using an ICC of 0.01, the study's sample size provides the statistical power necessary to detect a reduction of 6 percentage points in the occurrence of physical violence.
- **Stress of the pregnant woman**: Using an SD of 15.4 points (TMRI data), an ICC of 0.01 and the other parameters as defined above, we find that the study is powered to detect an impact on stress in women of 0.91 points (scale ranges from 0 to 40).
- **Stress of the husband of the pregnant woman:** Using an SD of 14.9 points (TMRI data), an ICC of 0.01 and the other parameters as defined above, we find that the study is powered to detect an impact on stress in women of 0.90 points (scale ranges from 0 to 40).
- **Maternal-fetal attachment**: Using a SD of 6 (39) and an ICC of 0.001, this sample size allows us to detect an effect of 0.83 for maternal-fetal attachment. This is equivalent to an effect size of 0.14.
- **Household food security**: Setting the prevalence of household food insecurity to 50% (the worst-case scenario in terms of power) and the ICC to 0.05 (equivalent to a design effect of 2.37), we can detect a change in the prevalence of food security of 10 percentage points. When the prevalence of food insecurity is 80%, the detectable difference is 8 to 9 percentage points.
- **Value of total household consumption**: Using an SD of 922 Tk (TMRI data), an ICC of 0.05 and the other parameters as defined above, we find that the study is powered to detect an impact on total per capita household food consumption of 193 Tk. Using an expected household size in the target population of 5 members, the effect we are able to detect is 965 Tk which is comparable to the size of the total transfer (1,000 Tk).
- **Value of household food consumption**: Using an SD of 621 Tk (TMRI data), an ICC of 0.05 and the other parameters as defined above, we find that the study are powered to detect an impact on total per capita household food consumption of 130 Tk. For an average 5-member household, the detectable difference are 650 Tk. We thus need to assume that at least 65% of the transfer are spent on food (top-up cash) or actually consumed by the household (food). This is a reasonable assumption, given that the BCC emphasizes the importance of using the transfers to improve the diet of the pregnant woman. In addition, TMRI households spent 62% of total expenditure on food.

Sample sizes for the primary outcomes measured immediately following birth are a bit smaller. We assume that 1.7% of the infants die in the first days after birth (33) and that an additional 5% of them are lost to follow-up, resulting in a sample size of around 1,076 infants on which birth outcome data can be collected. We further assume that children born the first 9 days of the 10-day break the enumerators take between survey waves (see section "fieldwork organization" below) is not measured within 72

hours after birth. This reduces the number of available newborns for birthweight assessment to 81/90 of the total number of live-born children or 968[6] newborns per arm.

Birthweight and low birthweight: Using a SD of birthweight in Bangladesh of 433 g (40), the study provides the necessary power to detect an impact on birthweight of 72 g. The available sample size would allow to detect an impact of 6 percentage points in the prevalence of low birthweight (the estimated current prevalence of low birthweight is 22.6% (40)). A systematic review and meta-analysis found that animal protein supplementation during pregnancy increased birthweight by 60 g (41). A systematic review and meta-analysis of balanced protein and energy supplements and food distribution programs found effects of 107 g and 46 g, respectively (42). We conclude that 72 g is a meaningful and feasible impact.

The sample size for the assessment when the child is 2 months (i.e., 61 days) of age was calculated as follows. We assume the same 1,076 children available for birth outcomes are available at 61 days postpartum, but given that the follow-up of 2-mo old infants ends by the end of wave 5, we estimate that around 1/4 of mothers and children do not contribute to the outcomes assessed at 61 days. We conducted the calculations with 17 observations per cluster or a total of 816 infants per arm (equivalent to 75.8%).

- **Prevalence of emotional violence towards the pregnant woman:** Assuming a prevalence of physical violence 40.0% (prevalence estimate from the Transfer Modality Research Initiative project or TMRI) and using an ICC of 0.01, the study's sample size provides the statistical power necessary to detect a change of 8 to 9 percentage points in the occurrence of physical violence (8 if the prevalence increases, 9 if the prevalence decreases).
- **Prevalence of controlling behaviors towards the pregnant woman:** Assuming a prevalence of controlling behaviors of 79.5% (prevalence estimate from the Transfer Modality Research Initiative project or TMRI) and using an ICC of 0.01, the study's sample size provides the statistical power necessary to detect a change of 7 to 8 percentage points in the occurrence of physical violence (7 if the prevalence increases, 8 if the prevalence decreases).
- **Prevalence of physical violence towards the pregnant woman**: Assuming a prevalence of physical violence of 19.9% (prevalence estimate from the Transfer Modality Research Initiative project or TMRI) and using an ICC of 0.01, the study's sample size provides the statistical power necessary to detect a change of 7 percentage points in the occurrence of physical violence (7 both in case the prevalence increases or decreases).
- **Stress of the pregnant woman**: Using an SD of 15.4 points (TMRI data), an ICC of 0.01 and the other parameters as defined above, we find that the study is powered to detect an impact on stress in women of 1.04 points (scale ranges from 0 to 40).
- **Maternal-infant attachment**: Using a SD of 4 (43), an ICC of 0.001, this sample size allows us to detect an effect of 0.65 for maternal-infant attachment (or an effect size of 0.16). Based on previous literature for the effects of behavioral interventions on maternal-infant attachment score, this effect size is appropriate for this outcome with infants under 12 months of age (44).

Notes on birth length and infant weight and length at 2 months (61 days) of age

[6] Due to the need to use integral numbers for the number of observations per cluster, we used 20 observations per cluster in the sample size calculations or 20 x 48 = 960 observations per arm. Enrollment is targeted to reach the 968 per arm which means that the actual observable difference is slightly smaller than what is shown here.

Sample size calculations for these outcomes reveal that the study is not powered to detect meaningful and feasible effects on these outcomes. The findings are summarized below:

- **Birth length:** using a SD of 2.12 mm (33), the study is powered to detect an effect on birth length of 3.5 mm. The length of newborns whose mothers received LNS during pregnancy was 2.4 mm higher than those whose mothers received IFA (45). A systematic review found a non-significant effect of 2.8 mm of balanced protein energy supplementation (only two studies were identified with a total of 67 participants) on birth length; a significant effect of 20 mm was found for food distribution programs (42). A study in Guatemala found that a family food ration of beans, rice, and oil and a ration of corn-soy blend targeted to the pregnant mother had a positive effect of 2.4 mm at 1 month postpartum (46). We conclude that the study is not sufficiently powered for birth length.
- **Infant length**: The detectable effect on length is 3.6 mm (using a SD of 2.02 mm at 61 days obtained from the WHO growth standards (47)). Evaluating whether this minimum detectable effect is meaningful and feasible is challenging, since evidence on the effect of pregnancy interventions on the length of infants 2 months is not available in the literature. We thus turn to the evidence on birth outcomes summarized above and conclude that the study is not adequately powered for this outcome.
- **Infant weight:** The mean SD for was obtained from the WHO growth standards (47). With a SD of 691 g at 61 days, the study has the necessary power to detect an impact on infant weight of 122 g which is larger than the impact we can reasonably expect to find based on previous studies.

3.4. Framework

We use an equivalence hypothesis testing framework. Primary analyses focus on the comparison between each of the treatment arms and the control arm, i.e., comparing arm 2 with arm 1 and comparing arm 3 with arm 1. Comparisons between arms 2 and 3 could be made as well, but the study is unlikely to be powered to detect meaningful effects.

3.5. Statistical interim analyses and stopping guidance

Not applicable.

3.6. Timing of final analysis

Analyses on a specific outcome start when data collection for that outcome has been finalized (see below) and data are made available by the data collection firm.

3.7. Timing of outcome assessments

Data are collected in the second and third trimester (i.e., 34 ± 1 wk) of pregnancy, within 72h after birth, and 2 months postpartum (i.e., 61 d ± 1 wk postpartum)(**Table 2**). Some outcomes are measured directly (e.g., prevalence of emotional violence); others require differencing assessments taken at two different points in time (e.g., gestational weekly weight gain).

Table 2: Timing of the measurement of the primary and secondary study outcomes

	Pregnancy trimester		Postpartum
	2nd	**3rd** (34 ± 1 wk)	**2 months#** (61 d ± 1 wk)
Primary outcomes			
Mother			
Gestational weekly weight gain 2nd to 3rd trimester	Yes	Yes	-
Dietary energy intake	Yes	Yes	-
Dietary protein intake	Yes	Yes	-
Mean probability of micronutrient adequacy	Yes	Yes	-
Length of gestation	Yes	-	Yes#
Preterm birth	-	-	Yes#
Prevalence of emotional violence towards the pregnant woman	Yes	Yes	Yes
Prevalence of controlling behaviors towards the pregnant woman	Yes	Yes	Yes
Prevalence of physical violence towards the pregnant woman	Yes	Yes	Yes
Stress	Yes	Yes	Yes
Maternal-fetal attachment	-	Yes	-
Maternal infant attachment	-	-	Yes
Husband			
Stress	Yes	Yes	
Infant			
Birthweight and low birthweight	-	-	Yes#
Household			
Household food security	Yes	Yes	-
Value of total household consumption	Yes	Yes	-
Value of household food consumption	Yes	Yes	-
Secondary outcomes			
Mother			
Gestational weekly weight gain up to 2nd trimester*	Yes		
Total gestational weight gain by the 3rd trimester*		Yes	
Use of iron-folic acid, calcium, vitamin B complex supplements[7], and multiple micronutrient supplements	Yes	Yes	-
Antenatal care utilization	Yes	Yes	-
Prenatal nutrition knowledge	Yes	Yes	-
Postnatal care utilization	-	-	Yes
Post-partum weight	-	-	Yes
Spousal relationship	Yes	Yes	Yes
Decision making	Yes	Yes	Yes
Mobility of the pregnant woman	Yes	Yes	Yes
Labor force participation	Yes	Yes	
Attitudes toward gender and intimate partner violence	Yes	Yes	Yes
Social capital	Yes	Yes	Yes
Agency	Yes	Yes	Yes

[7] The trial registration includes vitamin B complex supplementations. Since these vitamin supplements are not part of the Bangladesh National Strategy for Maternal Health, we dropped questions about vitamin B complex supplement use from the survey.

	Pregnancy trimester		Postpartum
	2nd	**3rd** (34 ± 1 wk)	**2 months#** (61 d ± 1 wk)
Prevalence of individual acts of intimate partner violence	Yes	Yes	Yes
Husband			
Prenatal nutrition knowledge	Yes	Yes	-
Infant			
Birth length	-	-	Yes#
Birth head circumference	-	-	Yes#
Head circumference	-	-	Yes
MUAC	-	-	Yes
Infant length and weight	-	-	Yes
Colostrum intake, (exclusive) breastfeeding practices, infant and young feeding practices	-	-	Yes
Neonatal mortality	-	-	Yes
Household			
Value of non-food consumption	Yes	Yes	-
Housing quality	Yes	No	Yes
Asset ownership	Yes	No	Yes
Income from remittances		Yes	
Savings	Yes	Yes	Yes
Loans	Yes	Yes	Yes
Participation in other social security programs		Yes	
Additional outcomes			
Mother			
Height&	Yes	-	-
MCB program enrollment	Yes	Yes	Yes
Receipt of base cash	Yes	Yes	Yes
Receipt of food basket and top-up cash	Yes	Yes	Yes
Participation in BCC sessions	Yes	Yes	Yes
Husband			
Height&	Yes	-	-
Weight&	Yes	-	-
Participation in BCC sessions	Yes	Yes	Yes
Household			
Economic shocks	-	-	Yes

These outcomes are assessed within 72 hours after birth.

* Maternal weight is also be measured during the census, i.e., before pregnancy

& Women's height and husband's height and weight are measured throughout the study period, i.e., when the team finds time to measure these outcomes.

Section 4 – Statistical principles

4.1. Level of statistical significance, confidence intervals to be reported, and adjustment for multiplicity

All statistical tests are two-sided. 95% confidence intervals are reported alongside exact p-values.

Multiplicity (i.e., carrying out many comparisons) has raised concerns about false-positive findings, which has led to recommendations to limit the number of comparisons made and/or to implement statistical adjustments. Multiplicity could arise in this trial in general for two reasons: having several outcomes of interest and having more than two study arms. We discuss both.

- When a set of hypotheses (each hypothesis about a different outcome) is tested simultaneously, the overall type I error rate (the probability of wrongly rejecting at least one null hypothesis), increases (18). Simultaneous hypothesis testing, however, assumes a universal null hypothesis, i.e., that the intervention has no effect for all the outcomes investigated versus the alternative hypothesis that the intervention had on at least one of these outcomes. Adjusting for multiplicity is only warranted if a set of hypotheses is tested simultaneously in this formal sense (18,19). This test of a universal null versus the stated alternative hypothesis is irrelevant to the scientific or evaluation questions being investigated for the primary outcomes in this study (18,20). When each hypothesis is limited to a single outcome, as is the case for the primary outcomes here, the probability of a false positive result remains the same irrespective of whether one or several comparisons are tested (18). Therefore, for all primary outcomes, no adjustment is made for multiplicity. Adjustments for multiplicity are made only in select cases when a universal null hypothesis is tested; in these cases, q-values are calculated to control the false discovery rate.
- Relative to the number of comparisons arising from multiple study arms, adjustments are not required if the number of comparisons for a given outcome is one less than the total number of arms (21). Since only two comparisons are planned for the primary analyses in this three-armed study, no adjustment is needed. The study is unlikely to be powered to detect meaningful effects between arms 2 and 3.

4.2. Adherence and protocol deviations

Adherence is defined as seeking enrollment in the program in the MCBP program in or before the 4th month of pregnancy, being enrolled subsequently, and receiving benefits as distributed by the program implementer until the end of pregnancy.

Adherence is presented in a table and in the paper narrative. In addition to full adherence, we present additional summary statistics to comprehensively describe program enrollment (i.e., being a beneficiary or not) and participation (i.e., receiving program benefits). These statistics, presented by trial arm, include (among others):

- Program enrollment:
    - the % of eligible women who never enrolled in the program
    - program enrollment by month of pregnancy
- Program participation among those enrolled in the program
    - % of women receiving 1, 2, … n transfers (base cash, top-up cash, food)
    - % of women participating in BCC (past month, over the course of the pregnancy)

A protocol deviation is defined as receiving the wrong treatment, e.g., receiving top-up cash when assigned to the food or base cash arm, receiving food when assigned to the top-up cash or base cash arm, and receiving the base cash only when assigned to the food or the top-up cash arm. Protocol deviations are be summarized by trial arm.

4.3. Analysis populations

The primary analysis is intent to treat (ITT). The ITT analysis includes all participants in the three arms who consented to participate in the study, enrolled in the study, met all program eligibility criteria (see section 5), and completed all baseline assessments. Treatment is defined as the treatment the study participants were randomized to receive. We also conduct a modified intent to treat analysis which only includes the study participants who enrolled in the MCBP (see section 6.2).

For our primary analyses, we do not conduct complete case analyses, as we attempt to use all available data from all participants. Robustness checks for select outcomes may restrict analysis to only the participants who have available data in all rounds.

Section 5 – Trial population

5.1. Screening data

A census was conducted to identify households with women who met the program's eligibility criteria (see Section 5.2). During the first, second, third, and fourth survey rounds, each household with a potentially eligible woman was visited to assess if the woman was pregnant in the second trimester. Pregnant women were invited to enroll in the study. Pregnant women are enrolled until the required round-specific sample size is reached.

5.2. Eligibility

The study enrolls women who meet the MCBP eligibility criteria[8]:

- 20-35 years of age;
- Pregnant with their first or second child[9] (in case of death of the first or second child during pregnancy or within two years of birth, the mother is eligible during her third pregnancy[10]);
- In possession of a valid NID (or a NID application acknowledgement).
- Lives in a household that meets one of the following poverty criteria:
    - The household head's main profession is rickshaw puller, van driver, day laborer, agricultural laborer, blacksmith, potter, washer man, fisher man, porter, or barber;
    - The household does not have a private latrine;
    - The household does not have an electricity connection while there is electricity in the village;
    - The household does not have a private tube well;
    - The household does not have an electric fan;
    - The household's place of residence has walls made of bamboo/jute sticks.

5.3. Recruitment

See section 5.1 for details. Eligible pregnant women were invited to enroll in the study and included after completing the informed consent process. Pregnant women are enrolled until the required sample size is reached.

A CONSORT flow diagram is used to summarize:

- The number of women who were assessed for eligibility (but not yet pregnant) at the time of the census and the number of clusters these women lived in;
- The number of women who were deemed eligible (but not yet pregnant) at the time of the census and the number of clusters these women lived in;
- The number of women who were deemed eligible (based on the census data) and pregnant at the time of the 1st, 2nd, 3rd, or 4th survey round and the number of clusters these women lived in;
- The number of women who were deemed eligible (based on the census data), pregnant at the time of the 1st, 2nd, 3rd, or 4th survey round, and enrolled in the study and the number of clusters these women lived in;
- The number of women who declined to participate in current and follow-up rounds and the number of women lost to follow-up (and the reasons why if available)
- The number of women included in the primary analysis
- The number of women excluded from the primary analysis (and the reasons why if available).

[8] Ineligible women who are inadvertently enrolled in the study are dropped from the analysis (see section 4.3).
[9] If she had twins the first time, this could be the third child.
[10] Even though the latest program directives no longer include this possibility, the program implementation follows this rule spelled out in an earlier version of the program directives (personal communication Mamunur Rashid, WFP; June 2023)

5.4. Withdrawal/follow-up

Withdrawals from the study are reported using two categories: # declined to participate in current and follow-up rounds; # lost contact/follow-up. Timing of withdrawals and loss to follow-up is presented in the CONSORT diagram, i.e. by study arm and survey round (i.e., 2nd trimester, 3rd trimester, birth, 2 months postpartum).

5.5. Baseline patient characteristics

Due to the nature of the study, we do not have true baseline data. We evaluate socio-demographic and other characteristics at study enrollment (2nd trimester of pregnancy) by which we describe our study sample (example of variables to be included in **Table 3**). We present these characteristics by intervention and control arms and report mean (and SD) and/or median (and the inter-quartile range) for continuous variables[11] and the N and percent for categorical and dichotomous variables. No significance testing or p-values are conducted or reported per CONSORT guidelines.

Table 3: Example of characteristics at enrollment to be summarized

Number of study subjects	
Socio-demographic	Age of the pregnant woman, household size, adult equivalents, marital status, educational attainment
Pregnancy	Gestational age (weeks) at enrollment, parity
Wealth	Housing quality, assets, food and non-food expenditure, household food insecurity
Anthropometry	Prepregnancy weight, height of the pregnant woman, BMI (kg/m2), % in each BMI category (underweight, normal, overweight, obese)

[11] Depending on the distribution of the continuous variable.

Section 6 – Analysis

6.1. Outcome definitions

The study includes three sets of outcomes: primary, secondary, and additional outcomes. The ***primary* study outcomes** are those outcomes that are predefined and for which the study is adequately powered to detect a feasible and meaningful effect (48).

The distribution of outcomes is inspected for normality and the variable transformed as necessary. More specific information on transformations is provided for key outcomes as needed.

- **Pregnant mother**
  - **Gestational weekly weight gain 2nd to 3rd trimester:** Women's weekly weight gain (kg/wk) is calculated by differencing women's weight obtained in the 2nd and 3rd trimester and dividing by the number of weeks between both measurements.
  - **Dietary energy intake and proportion of women below 85% of the estimate energy requirement (EER)**: Dietary energy intake (kcal/day) is assessed using a 24hr recall in the second and third trimester (34). EER is calculated using each woman's basal metabolic rate (estimated from the woman's age, gender, and current weight), level of physical activity, and pregnancy trimester. Factors of 1.4 for low, 1.7 for moderate, and 2.0 for high physical activity is used (Food and Agricultural Organization, 2001). Additional energy requirement for the second and third trimester of pregnancy (340-350 kcal/d and 452-500 kcal/d, respectively, depending on which reference is used (35,36)) is added to account for gestational weight gain and increases in basal metabolic rate. Impact is assessed in the 2nd and 3rd trimester. Given continued advances in the analysis of 24h dietary recall data, the final analytic strategy to calculate the adequacy for each nutrient is based on recommended best practice at the time of the analyses.
  - **Dietary protein intake:** The 24hr recall data is used to assess women's protein intake (g/day). Impact is assessed in the 2nd and 3rd trimester. Given continued advances in the analysis of 24h dietary recall data, the final analytic strategy to calculate the adequacy for each nutrient is based on recommended best practice at the time of the analyses.
  - **Mean probability of micronutrient adequacy:** the mean of the probabilities of adequacy for 11 key micronutrients (iron, calcium, zinc, vitamin A, thiamin, riboflavin, niacin, vitamin B-6, vitamin B-12, vitamin C, and folate) is calculated using the 24hr recall data. Given continued advances in this field, the final analytic strategy to calculate the adequacy for each nutrient is based on recommended best practice at the time of the analyses. For now, we anticipate using the methods as used by Nguyen and colleagues (Nguyen et al., 2018); and the WHO's estimated average requirements (EAR) to calculate the probability of adequacy for each micronutrient (World Health Organization, 2004; World Health Organization & Food and Agriculture Organization of the United Nations, 2004). Usual intakes are estimated using the data from a second 24hr recall collected on a non-consecutive day for a subsample of pregnant women. A power transformation is used to obtain symmetrical distributions for each nutrient; and a measurement-error model is applied to adjust for the within person variance of nutrient intake (34,52). The estimated usual intake is used to calculate the probability that the usual intake was above the EAR during pregnancy (53). Impact is assessed in the 2nd and 3rd trimester.
  - **Length of gestation:** The difference between the date of birth and the first day of the last menstrual period (assessed in the first survey) is used to calculate length of gestation (days).
  - **Preterm birth:** Using the length of gestation, we determine the proportion of children born before 37 weeks (or 259 days) of gestation.

- **Prevalence of emotional violence towards the pregnant woman/mother of the newborn child**: Prevalence of any emotional intimate partner violence in the past 6 months measured using the WHO Violence Against Women instrument (54). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months (i.e., 61 days) postpartum.
- **Prevalence of controlling behaviors towards the pregnant woman/mother of the newborn child**: Prevalence of any controlling behaviors in the past 6 months measured using the WHO Violence Against Women instrument (54). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months (i.e., 61 days) postpartum.
- **Prevalence of physical violence towards the pregnant woman/mother of the newborn child**: Prevalence of any physical intimate partner violence in the past 6 months, measured using the WHO Violence Against Women instrument (54). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months (i.e., 61 days) postpartum.
- **Stress of the pregnant woman/mother of the newborn child**: Measured using women's Perceived Stress Scale score (55). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum.
- **Maternal-fetal attachment:** Maternal attachment to their unborn child is assessed during their 3rd trimester of pregnancy using the Prenatal Attachment Inventory (56). In the Bangla translation of the Prenatal Attachment Inventory, all items in the scale were rephrased as questions instead of statements to make it easier for the interviewer to ask and to enhance the respondents' understanding.
- **Maternal-infant attachment:** Maternal attachment to their infant is assessed at 2 months (i.e., 61 days) postpartum using the Postpartum Bonding Questionnaire (57). The items in the Bangla-version of the scale were rephrased as questions instead of statements to simplify the interview process. Additionally, 2 of the 25 items in the original scale were removed. These 2 items (i.e., "I have done harmful things to my baby" and "I feel like hurting my baby") made up the subscale on Risk of Abuse and had low sensitivity (57). The sum of all items on the full scale, as well as each of the three subscales (Bonding, mother-infant relationship, and infant-focused anxiety), is used as outcome indicators.

- **Husband/partner of the pregnant woman**:
  - **Stress of the husband of the pregnant woman**: Measured using men's Perceived Stress Scale score (55). The second- and third-trimester assessment is used to assess impact.
- **Infant**
  - **Birthweight**: Birthweight (grams) is assessed within 72 hours after delivery to avoid the influence of transitory neonatal weight loss that typically happens during the first days of life.
  - **Low birthweight**: Low birthweight is defined as a birth weight below 2500 grams.
- **Household**
  - **Household food security**: The Household Food Insecurity Access Scale is used to measure household food security status (58). The 2nd and 3rd trimester assessment is used to assess impact. Impact is assessed on the continuous scale (ranging from 0 to 27) and the categorical scale (food secure, mildly food insecure, moderately food insecure, or severely food insecure). The actual household groupings based on the categorical outcome depend on the population distribution.
  - **Value of total household consumption**: Aggregate value of household food and non-food consumption expenditures (Taka). This is a continuous measure calculated from household survey responses on consumption behavior, using the methodology and questionnaire modules described by Deaton and Zaidi (59). The 2nd and 3rd trimester assessment is used to assess impact. We assess inflation during the study period and deflate values as needed to be equivalent to the first wave in 2023. The distribution of this outcome is inspected for normality and the variable is transformed as necessary.

- **Value of household food consumption**: Value of household food consumption expenditures (Taka). This is a continuous measure calculated from household survey responses on consumption behavior, using the methodology and questionnaire modules described by Deaton and Zaidi (59). The 2nd and 3rd trimester assessment is used to assess impact. We assess inflation during the study period and deflate values as needed to be equivalent to the first wave in 2023. The distribution of this outcome is inspected for normality and the variable is transformed as necessary.

The ***secondary*** **study outcomes** are those outcomes on which the intervention is expected to have an impact but for which the study is unlikely to be adequately powered or for which statistical power cannot easily be calculated.

- **Pregnant mother**
  - **Gestational weekly weight gain up to 2nd trimester:** Women's weekly weight gain (kg/wk) is calculated by differencing women's weight obtained at the census (before pregnancy) and 2nd trimester and dividing by gestational age in weeks.
  - **Total gestational weight by the third trimester:** Total weight gain (kg) is calculated by calculating the difference between women's third trimester weight and the weight obtained at the time of the census.
  - **Use of iron-folic acid, and calcium, and vitamin B complex, and multiple micronutrient supplements[12]:** Women are asked to report on their use of these supplements during pregnancy. The 2nd and 3rd trimester assessment is used to assess impact. We report on % of women who ever consumed these supplements during this pregnancy, the % who consumed the supplement in the past 7 d, and the frequency of consumption in the past 7d.
  - **Antenatal care utilization**: Antenatal care utilization is assessed using a variety of measures based on current recommendations in Bangladesh and the new WHO guidelines (60). The 2nd and 3rd trimester and postpartum[13] assessment is used to assess impact.
  - **Prenatal nutrition knowledge**: Women's knowledge related to nutrition is assessed using a set of questions developed for this study. The 2nd and 3rd trimester assessment is used to assess impact. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
  - **Postnatal care utilization**: Postnatal care utilization is assessed using a variety of measures based on current recommendations in Bangladesh and WHO guidelines. The 2-month postpartum assessment is used to assess impact.
  - **Postpartum weight**: We measure women's weight (kg) at the 2-month postpartum visit.
  - **Spousal relationship between the pregnant woman/mother of the newborn child and her husband**: Measured using women's spousal relationship modules developed for the survey. Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
  - **Decision making by the pregnant woman/mother of the newborn child**: Measured using women's decision-making modules adapted from pro-WEAI (61). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum. The construction of the outcome

[12] The trial registration includes vitamin B complex supplementations. Since these vitamin supplements are not part of the Bangladesh National Strategy for Maternal Health, we dropped questions about vitamin B complex supplement use from the survey.

[13] The trial registration incorrectly limits this outcome to the 2nd and 3rd trimester observations. Data were collected on this outcome postpartum as well and this information is used to assess impact on this outcome for the entire pregnancy.

measure(s) is based on the distribution of and the association between the variables on which data were collected.

- **Mobility of the pregnant woman**: Measured using mobility modules adapted from pro-WEAI (61). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
- **Labor force participation of the pregnant woman**: Measured using women's labor modules developed for the survey. The 2nd and 3rd trimester assessment is used to assess impact. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
- **Attitudes toward gender and intimate partner violence of the pregnant woman**: Measured using women's attitudes toward gender and intimate partner violence modules developed for the survey. Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
- **Social capital of the pregnant woman**: Measured using modules on women's social interactions and perceptions of neighborhood responses to intimate partner violence, developed for the survey. Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
- **Agency of the pregnant woman**: Measured using women's responses on locus of control (62). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
- **Prevalence of individual acts of intimate partner violence towards the pregnant woman/mother of the newborn child**: Prevalence of individual acts of intimate partner violence in the past 6 months measured using the WHO Violence Against Women instrument (54). Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum.
- **Institutional delivery**: Delivery at a health facility as reported by the mother.
- **Skilled attendance at birth:** Whether a skilled birth attendant attended the birth of the child as reported by the mother.
- **Caesarian section:** Women are asked if a c-section was used to deliver the child.

- **Husband/partner of the pregnant woman**:
  - **Prenatal nutrition knowledge of the husband of the pregnant woman**: Using a similar approach used for pregnant women, husband's nutrition knowledge. The 2nd and 3rd trimester assessment is used to assess impact. The outcome is constructed the same way as the nutrition knowledge outcome for pregnant woman.
  - **Weight:** We measure the husband's weight (kg) at each visit. Impact is assessed in the 2nd and 3rd trimester of pregnancy and at 2 months postpartum.
- **Infant**:
  - **Birth length**: The length of the newborn (cm) is assessed at the time of the birthweight assessment.
  - **Infant length and weight:** Length (cm) and weight (kg) of infants 2 months (61 days) of age is measured.
  - **Colostrum intake, time between birth and putting infant to breast, (exclusive) breastfeeding after birth:** outcomes are defined as in the WHO guidance on indicators for assessing infant and young child feeding practices (63). The assessment at birth is used.

- **(Exclusive) breastfeeding and infant feeding practices at 2 months:** outcomes are defined as in the WHO guidance on indicators for assessing infant and young child feeding practices (63). Impact is assessed at 2 months postpartum.
- **Neonatal and infant mortality**: neonatal (within the first 1 month after birth) and infant (within the first 2 months after birth) mortality is assessed through parental recall at the 2-month postpartum visit.

- **Household**:
  - **Value of household non-food consumption**: Value of household non-food consumption expenditures. This is a continuous measure calculated from household survey responses on consumption behavior, using the methodology and questionnaire modules described by Deaton and Zaidi (59). The 2nd and 3rd trimester assessment is used to assess impact. We assess inflation during the study period and deflate values as needed to be equivalent to the first wave in 2023. The distribution of this outcome is inspected for normality and the variable is transformed as necessary.
  - **Housing quality**: Quality of housing stock, measured using a module developed for the study context and collected through household enumerator observation during interviews. The 2nd trimester and 2-months postpartum assessment is used to assess impact. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
  - **Household asset ownership**: Index of household asset ownership, measured using a module developed for the study context. The 2nd trimester and 2-months postpartum assessment is used to assess impact. The construction of the outcome measure(s) is based on the distribution of and the association between the variables on which data were collected.
  - **Household income from remittances:** Measured with a module on income (Taka) from remittances received from relatives outside and within the country. We calculate the sum of all remittances received. The 3rd trimester assessment is used to assess impact.
  - **Household savings:** A continuous measure of total household savings, calculated using reported savings of each household member 15 years of age or older. The 2nd and 3rd trimester and 2-month postpartum assessments are used to assess impact. We calculate the sum of all household savings.
  - **Household loans**: A continuous measure on household loans, calculated using reported amounts of loans taken and outstanding amounts owed for each household member 15 years of age or older. The 2nd and 3rd trimester and 2-month postpartum assessments is used to assess impact. We calculate the sum of all household loans.

In addition to the primary and secondary study outcomes, the study collects rich data on a set of ***additional* study outcomes**. These include variables describing program participation (such as attendance at BCC sessions) and variables that can reduce residual noise in the models used to estimate impact (such as maternal height and shocks). A non-exhaustive list of examples is provided below.

- **Pregnant mother**
  - **Height**: height (cm) of the pregnant woman
  - **MCBP program enrollment**: whether the mother (and her newborn child) are enrolled in MCBP.
  - **Receipt of base cash:** whether the mother has been receiving the base cash.
  - **Receipt of food basket and top-up cash:** whether the mother has been receiving the food basket or top-up cash.
  - **Participation in BCC sessions:** whether the mother (and other household members) have been been participating in the BCC sessions.
- **Husband**

- **Height**: height (cm) of the husband
- **Household**
  - **Economic shocks:** the occurrence of shocks affecting the household in the months preceding the survey.

6.2. Analysis method

Intent to treat analyses
Intent to treat analyses are used to assess whether the intervention resulted in changes in primary and secondary outcomes.

The primary intent to treat analyses uses mixed-effects models.

For continuous outcomes available at only one point in time (Y_{ij}, such as gestational weight gain from the 2nd to the 3rd trimester and length of gestation), we use a linear mixed-effects model with study cluster (u_j) and individual (ε_{ij}) as random effects assuming an exchangeable error structure, and treatment arm (A_j) and covariates (X_{ij}, see below) as fixed effects.

$$Y_{ij} = A_j\beta + X_{ij}\gamma + u_j + \varepsilon_{ij} \qquad (1)$$

For continuous outcomes available at multiple points in time (Y_{ijt}, such as dietary intake), we use linear mixed-effects models that use all of the longitudinal data with study cluster (u_j), individual (α_{ij}), and time (ε_{ijt}) as a random effects assuming an exchangeable error structure, and treatment arm (A_j), survey round (T_t), the interaction between study arm and survey round ($A_j \cdot T_t$), and covariates (X_{ijt}, see below) as fixed effects.

$$Y_{ijt} = A_j\beta + X_{ijt}\gamma + T_t\delta + A_j \cdot T_t\theta + u_j + \alpha_{ij} + \varepsilon_{ijt} \quad (2)$$

The same linear mixed models are used for dichotomous outcomes (such as preterm birth and prevalence of emotional violence), but we use robust standard errors to account for heteroscedascity.

For select outcomes assessed at multiple points in time (i.e., emotional violence towards the pregnant woman/mother of the newborn child; controlling behaviors towards the pregnant woman/mother of the newborn child; and physical violence towards the pregnant woman/mother of the newborn child) we assess the impact on the occurrence at any point in time (i.e., at either the 2nd trimester, 3rd trimester, or 2 months postpartum survey) using a model as shown in (1).

Even though randomization should yield equivalence between arms on covariates, we include covariates or make other adjustments (such as re-weighting) in case of non-equivalence of enrollment measures or due to differential attrition. In addition, we include covariates (tentative list in Table 4) that are orthogonal to the randomized treatment assignment and are known or suspected predictors of the outcome of interest (such as, maternal age and education) to reduce residual noise (i.e., increase statistical power). A tentative list of covariates is presented in Table 4. Decisions on which covariates to ultimately include is based on factors such as the extent to which potential covariates are correlated with outcomes, the share of observations with missing values of potential covariates, and the extent to which potential covariates have realistic distributions.

Table 4: Tentative list of covariates to be included in regression models

Outcome	Covariates
All outcomes	Socio-economic status (using census data), upazila, woman's educational attainment, woman's age, season, having an NID
Dietary outcomes	Additional covariates: Interviewer, day of week
Prevalence of emotional violence, controlling behaviors, physical violence, stress	Additional covariates: partner's age, partner's educational attainment, woman's gestational age
Attachment	Additional covariates: -
Birth outcomes	Additional covariates: Maternal height, maternal BMI at census, primiparity, child sex
Household food security, value of household consumption	Household adult equivalents

Distributional assumptions for outcome variables is checked using descriptive statistics. Transformations are used for continuous outcome variables if needed to obtain more symmetric distributions.

In addition to the models described above, we conduct robustness checks by estimating fixed effects models.

The intent to treat analyses do not include any subgroup analyses (e.g., by maternal age, season, etc.)

Modified intent to treat analysis

Based on field reports, a considerable proportion of women who are eligible for MCBP and included in our study may not have enrolled in MCBP. Consequently, if the addition of top-up cash or food has effects on the study outcomes over and above the base MCBP, these effects may not be detected in the intent to treat analysis, as the impact in those enrolled is diluted by the absence of impact in those not in the program. If we find that enrollment in the program is indeed low, we will consider conducting modified intent to treat analysis, to assess if the program had an impact within the subgroup who enrolled. The exact method used depends on patterns in the data.

It is possible that some of the difference in enrollment across arms could be explained by differences in observables across arms, but we do not yet know. If there appears to be selection bias in enrollment that differs by treatment arm, we will use quasi-experimental analytical methods. For example, inclusion of additional covariates, propensity score methods, or matching methods may be appropriate if women's observable pre-pregnancy characteristics measured in the census appear to predict well whether they enroll; if the distributions of these observable characteristics show meaningful overlap across arms; and if empirical patterns do not suggest a strong role for unobserved factors that both predict enrollment and differ systematically by arm.

Exploratory analysis

Exploratory analyses are conducted to assess the heterogeneity of impact, i.e. to evaluate whether the impact was different in specific subgroups (such as women in more vs. less disadvantaged households, women with lower vs. higher levels of education, women of younger vs. older age, women who co-reside with mothers-in-law vs. women who do not co-reside with mothers-in-law at 2nd trimester) or by season. To assess the heterogeneity of impact, we stratify the analyses on the possible effect modifier. The analyses are considered exploratory as they are not adequately powered.

6.3. Missing data

Approaches for handling missing data depend on the outcome or covariate.

For instances when a limited amount of information is missing that is needed to construct an outcome variable (e.g., the price of a consumed food which is needed to construct total household expenditure), we impute the missing information using single imputation drawing on the non-missing data.

When outcomes (such as birthweight) or covariates are missing, decisions on whether and how to impute missing values is based on assessing patterns in the data, such as the share of missing values, the expected reason for missing values, and whether there is sufficient other information to impute missing values well. Outcomes assessed at multiple times that are missing are not imputed because the mixed model handles these missing values assuming that they are missing at random.

If an outcome has a large share of missing values, robustness checks may be conducted to assess if impact estimates may suffer from selection bias. This may include comparing the characteristics of households or individuals for which the outcome is missing with those for whom the outcome is not missing and the calculation of Lee Bounds.

6.4. Additional analyses

Exploratory additional analyses may be conducted to quasi-experimentally estimate the absolute impact of the base MCBP program, relative to no MCBP. Whether these analyses are feasible depends on patterns in the data. For example, we will assess the share of women eligible for MCBP who do not enroll, the observable characteristics measured in the census that predict enrollment, and the overlap of these observable characteristics between women enrolled in the base MCBP program and women who do not enroll in MCBP.

6.5. Harms

The study team used several strategies to insure the highest levels of data safety. Enumerators were extensively trained on research ethics and the importance of confidentiality. Data were collected on password protected tablets. Data were uploaded to a secure server which could only be accessed by the members of the study team. No PII data have or will be shared with external investigators.

6.6. Statistical software

Statistical analyses are conducted using Stata and R.

References

1. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, de Onis M, Ezzati M, Grantham-McGregor S, Katz J, Martorell R, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet. 2013;382:427–51.

2. Arsenault JE, Moursi M, Sraboni E, Ahmed AU. Intra-household allocation of food and nutrient intakes in a nationally representative sample of Bangladesh (draft).

3. ICDDR,B, UNICEF, GAIN I. National Micronutrient Survey in Bangladesh in 2011-12. Dhaka, Bangladesh: Centre for Nutrition and Food Security, ICDDR,B; 2013.

4. Arsenault JE, Moursi M, Sraboni E, Ahmed AU. Intra-household allocation of food and nutrient intakes ina nationally representative sample of Bangladesh (draft).

5. National Nutrition Services (NNS) of Institute of Public Health Nutrition (IPHN). National Low Birth Weight Survey 2015. Bangladesh. Dhaka, Bangladesh; 2017.

6. Nguyen PH, Sanghvi T, Kim S, Mahmud Z, LM T, Shabnam S, Aktar B, Afsana K, Frongillo E, Ruel M, et al. Integrating nutrition interventions into an existing maternal, neonatal and child health platform increased maternal dietary diversity and micronutrients intake and exclusive breastfeeding practices in Bangladesh: Results of a cluster-randomized program e. Journal of Nutrition.

7. NIPORT. Bangladesh Demographic and Health Survey 2014. Dhaka, Bangladesh and Calverton, MD: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International; 2015.

8. Ahmed AU, Ahmad K, Chou V, Hernandez R, Menon P, Naeem F, Naher F, Quabili W, Sraboni E, Yu B. The Status of Food Security in the Feed the Future Zone and Other Regions of Bangladesh: Results from the 2011–2012 Bangladesh Integrated Household Survey. International Food Policy Research Institute (IFPRI), Washington, D.C.; 2013.

9. Owais A, Faruque ASG, Das SK, Ahmed S, Rahman S, Stein AD. Maternal and Antenatal Risk Factors for Stillbirths and Neonatal Mortality in Rural Bangladesh: A Case-Control Study. Moormann AM, editor. PLoS One. 2013;8:e80164.

10. Mercer A, Haseen F, Huq NL, Uddin N, Hossain Khan M, Larson CP. Risk factors for neonatal mortality in rural areas of Bangladesh served by a large NGO programme. Health Policy Plan. 2006;21:432–43.

11. NHSR. Bangladesh. National Newborn Health Situation Analysis Report. Dhaka, Bangladesh; 2014.

12. Lee ACC, Katz J, Blencowe H, Cousens S, Kozuki N, Vogel JP, Adair L, Baqui AH, Bhutta ZA, Caulfield LE, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. Lancet Glob Health. 2013;1:e26-36.

13. Ruel MT, Alderman H. Nutrition-sensitive interventions and programmes: how can they help to accelerate progress in improving maternal and child nutrition? Lancet. 2013 Aug.

14. Leroy JL, Koch B, Roy S, Gilligan D, Ruel M. Social Assistance Programs and Birth Outcomes: A Systematic Review and Assessment of Nutrition and Health Pathways. J Nutr. 2021;151:3841–55.

15. World Health Organization. WHO recommendations on antenatal care for a positive pregnancy experience. Geneva, Switzerland: World Health Organization; 2016.

16. Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. Int J Epidemiol. MRC Tropical Epidemiology Group, London School of Hygiene and Tropical Medicine, UK; 1999;28:319–26.

17. Hemming K, Marsh J. A menu-driven facility for sample-size calculations in cluster randomized controlled trials. Stata Journal. 2013;13:114–35.

18. Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. The Lancet [Internet]. 2005;365:1591–5. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673605664616

19. U.S. Department of Health and Human Services - Food and Drug Administration. Guidance for Industry: Multiple Endpoints in Clinical Trials - Draft guidance. Rockville, Maryland, USA; 2017.

20. Perneger T V. What's wrong with Bonferroni adjustments. British Medical Journal. BMJ Publishing Group; 1998. p. 1236–8.

21. Steel R, Torrie J. Principles and Procedures of Statistics: A Biometrical Approach. Second Edi. New York: McGraw-Hill Book Company; 1980.

22. Adu-Afarwuah S, Lartey A, Okronipa H, Ashorn P, Ashorn U, Zeilani M, Arimond M, Vosti SA, Dewey KG. Maternal Supplementation with Small-Quantity Lipid-Based Nutrient Supplements Compared with Multiple Micronutrients, but Not with Iron and Folic Acid, Reduces the Prevalence of Low Gestational Weight Gain in Semi-Urban Ghana: A Randomized Controlled Trial. J Nutr. 2017;147:697–705.

23. Matias SL, Mridha MK, Paul RR, Hussain S, Vosti SA, Arnold CD, Dewey KG. Prenatal Lipid-Based Nutrient Supplements Affect Maternal Anthropometric Indicators Only in Certain Subgroups of Rural Bangladeshi Women. J Nutr. 2016;146:1775–82.

24. Hambidge KM, Westcott JE, Garcés A, Figueroa L, Goudar SS, Dhaded SM, Pasha O, Ali SA, Tshefu A, Lokangaka A, et al. A multicountry randomized controlled trial of comprehensive maternal nutrition supplementation initiated before conception: the Women First trial. Am J Clin Nutr. 2019;109:457–69.

25. Ota E, Hori H, Mori R, Tobe-Gai R, Farrar D. Antenatal dietary education and supplementation to increase energy and protein intake. Ota E, editor. Cochrane Database Syst Rev. Chichester, UK, UK: John Wiley & Sons, Ltd; 2015;CD000032.

26. Institute of Medicine (US) and National Research Council (US) Committee to Reexamine IOM Pregnancy Weight Guidelines. Weight Gain During Pregnancy: Reexamining the Guidelines. Rasmussen KM, Yaktine AL, editors. Washington (DC), US: National Academies Press; 2009.

27. Kac G, Arnold CD, Matias SL, Mridha MK, Dewey KG. Gestational weight gain and newborn anthropometric outcomes in rural <scp>Bangladesh</scp>. Matern Child Nutr. 2019;15.

28. Gernand AD, Christian P, Paul RR, Shaikh S, Labrique AB, Schulze KJ, Shamim AA, West KP. Maternal Weight and Body Composition during Pregnancy Are Associated with Placental and Birth Weight in Rural Bangladesh. J Nutr. 2012;142:2010–6.

29. Hasan SMT, Rahman S, Locks LM, Rahman M, Hore SK, Saqeeb KN, Khan MdA, Ahmed T. Magnitude and determinants of inadequate third-trimester weight gain in rural Bangladesh. Drehmer M, editor. PLoS One. 2018;13:e0196190.

30. Alam DS, van Raaij JMA, Hautvast JGAJ, Yunus M, Fuchs GJ. Energy stress during pregnancy and lactation: consequences for maternal nutrition in rural Bangladesh. Eur J Clin Nutr. 2003;57:151–6.

31. Christian P, Khatry SK, Katz J, Pradhan EK, LeClerq SC, Shrestha SR, Adhikari RK, Sommer A, West Jr. KP. Effects of alternative maternal micronutrient supplements on low birth weight in rural Nepal: double blind randomised community trial. Bmj. Division of Human Nutrition, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA. pchristi@jhsph.edu; 2003;326:571.

32. West KP, Shamim AA, Mehra S, Labrique AB, Ali H, Shaikh S, Klemm RDW, Wu LS-F, Mitra M, Haque R, et al. Effect of Maternal Multiple Micronutrient vs Iron–Folic Acid Supplementation on Infant Mortality and Adverse Birth Outcomes in Rural Bangladesh. JAMA [Internet]. 2014;312:2649. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2014.16819

33. Mridha MK, Matias SL, Chaparro CM, Paul RR, Hussain S, Vosti SA, Harding KL, Cummins JR, Day LT, Saha SL, et al. Lipid-based nutrient supplements for pregnant women reduce newborn stunting in a cluster-randomized controlled effectiveness trial in Bangladesh. Am J Clin Nutr [Internet]. 2016;103:236–49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26607935

34. Nguyen PH, Huybregts L, Sanghvi TG, Tran LM, Frongillo EA, Menon P, Ruel MT. Dietary Diversity Predicts the Adequacy of Micronutrient Intake in Pregnant Adolescent Girls and Women in Bangladesh, but Use of the 5-Group Cutoff Poorly Identifies Individuals with Inadequate Intake. J Nutr [Internet]. 2018;148:790–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30053285

35. Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. Washington, D.C.: National Academies Press; 2006.

36. Butte NF, Wong WW, Treuth MS, Ellis KJ, O'Brian Smith E. Energy requirements during pregnancy based on total energy expenditure and energy deposition. Am J Clin Nutr. 2004;79:1078–87.

37. Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. Washington, D.C.: National Academies Press; 2006.

38. Gernand AD, Paul RR, Ullah B, Taher MA, Witter FR, Wu L, Labrique AB, West KP, Christian P. A home calendar and recall method of last menstrual period for estimating gestational age in rural Bangladesh: a validation study. J Health Popul Nutr. 2016;35:34.

39. Edhborg M, Nasreen H-E, Kabir ZN. Impact of postpartum depressive and anxiety symptoms on mothers' emotional tie to their infants 2–3 months postpartum: a population-based study from rural Bangladesh. Arch Womens Ment Health. 2011;14:307–16.

40. National Nutrition Services (NNS) of Institute of Public Health Nutrition (IPHN). National Low Birth Weight Survey 2015. Bangladesh. Dhaka, Bangladesh; 2017.

41. Pimpin L, Kranz S, Liu E, Shulkin M, Karageorgou D, Miller V, Fawzi W, Duggan C, Webb P, Mozaffarian D. Effects of animal protein supplementation of mothers, preterm infants, and term infants on growth outcomes in childhood: a systematic review and meta-analysis of randomized trials. Am J Clin Nutr. 2019;110:410–29.

42. Lassi ZS, Padhani ZA, Rabbani A, Rind F, Salam RA, Das JK, Bhutta ZA. Impact of Dietary Interventions during Pregnancy on Maternal, Neonatal, and Child Outcomes in Low- and Middle-Income Countries. Nutrients. 2020;12:531.

43. Wulff V, Hepp P, Wolf OT, Fehm T, Schaal NK. The influence of maternal singing on well-being, postpartum depression and bonding – a randomised, controlled trial. BMC Pregnancy Childbirth. 2021;21:501.

44. van Lieshout RJ, Layton H, Savoy CD, Brown JSL, Ferro MA, Streiner DL, Bieling PJ, Feller A, Hanna S. Effect of Online 1-Day Cognitive Behavioral Therapy–Based Workshops Plus Usual Care vs Usual Care Alone for Postpartum Depression. JAMA Psychiatry. 2021;78:1200.

45. Das JK, Hoodbhoy Z, Salam RA, Bhutta AZ, Valenzuela-Rubio NG, Weise Prinzo Z, Bhutta ZA. Lipid-based nutrient supplements for maternal, birth, and infant developmental outcomes. Cochrane Database of Systematic Reviews. 2018;2018.

46. Olney DK, Leroy J, Bliznashka L, Ruel MT. PROCOMIDA, a Food-Assisted Maternal and Child Health and Nutrition Program, Reduces Child Stunting in Guatemala: A Cluster-Randomized Controlled Intervention Trial. J Nutr. 2018;148:1493–505.

47. WHO multicentre growth reference study group. WHO Child Growth Standards. Acta Paediatr Suppl. 2006;95:1–104.

48. Leroy JL, Frongillo EA, Kase BE, Alonso S, Chen M, Dohoo I, Huybregts L, Kadiyala S, Saville NM. Strengthening causal inference from randomised controlled trials of complex interventions. BMJ Glob Health. 2022;7:e008597.

49. Food and Agricutural Organization. Human energy requirements: Report of a Joint FAO/WHO/UNU Expert Consultation. FAO Food and Nutrition Technical Report Series. Rome; 2004 Oct.

50. World Health Orginazation. Vitamin and mineral requirements in human nutrition. Geneva; 2004.

51. World Health Orginazation, Food and Agriculture Organization of the United Nations. Human vitamin and mineral requirements. Report of a joint FAO and WHO expert consultation held in Bangkok, Thailand, 21–30 September 1998. Rome; 2004.

52. Joseph ML, Carriquiry A. A measurement error approach to assess the association between dietary diversity, nutrient intake, and mean probability of adequacy. J Nutr. 2010;140:2094S-101S.

53. Subcommittees on Interpretation and Uses of Dietary Reference Intakes and the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes Food and Nutrition. Dietary reference intakes. Applications in dietary assessment. Washington, D.C.: National Academy Press; 2000.

54. Ellsberg M, Heise L. Researching Violence Against Women: A Practical Guide for Researchers and Activists. Washington DC; 2005.

55. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. J Health Soc Behav. 1983;24:385.

56. Muller ME, Mercer RT. Development of the Prenatal Attachment Inventory. West J Nurs Res. 1993;15:199–215.

57. Brockington IF, Fraser C, Wilson D. The Postpartum Bonding Questionnaire: a validation. Arch Womens Ment Health. 2006;9:233–42.

58. Coates J, Swindale A, Bilinsky P. Household Food Insecurity Access Scale (HFIAS) for measurement of food access: Indicator Guide (v.3). Washington D.C.: Food and Nutrition Technical Assistance Project (FANTA)/Academy for Educational Development; 2007.

59. Deaton A, Zaidi S. Guidelines for constructing consumption aggregates for welfare analysis. 2002.

60. WHO. WHO recommendations on antenatal care for a positive pregnancy experience. Geneva, Switzerland: World Health Organization; 2016.

61. Malapit H, Quisumbing A, Meinzen-Dick R, Seymour G, Martinez EM, Heckert J, Rubin D, Vaz A, Yount KM. Development of the project-level Women's Empowerment in Agriculture Index (pro-WEAI). World Dev. 2019;122:675–92.

62. Rotter JB. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs: General and Applied. 1966;80:1–28.

63. World Health Organization, United Nations Children's Fund (UNICEF). Indicators for assessing infant and young child feeding practices: definitions and measurement methods. World Health Organization; 2021. viii, 110 p. p.

64. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318:2337.

65. Merrill RD, Shamim AA, Ali H, Labrique AB, Schulze K, Christian P, West KP. High prevalence of anemia with lack of iron deficiency among women in rural Bangladesh: a role for thalassemia and iron in groundwater. Asia Pac J Clin Nutr. 2012;21:416–24.