Author’s response to reviews

Title: Iron-fortified lentils to improve iron (Fe) status among adolescent girls in Bangladesh- study protocol for a double-blind community-based randomized control trial

Authors:

Fakir Yunus (fakir.yunus@usask.ca)
Chowdhury Jalal (cjalal@NUTRITIONINTL.ORG)
Kaosar Afsana (afsana.k@brac.net)
Rajib Podder (rajib.podder@usask.ca)
Albert Vandenberg (bert.vandenberg@usask.ca)
Diane DellaValle (ddellavalle@marywood.edu)

Version: 1 Date: 21 Dec 2018

Author’s response to reviews:

RESPONSE TO REVIEWER COMMENTS

TRLS-D-18-00783

Iron-fortified lentils to improve iron (Fe) status among adolescent girls in Bangladesh- study protocol for a double-blind community based randomized control trial

Fakir Md Yunus, MBBS., MPH., MSc.; Chowdhury Jalal, PhD; Kaosar Afsana, PhD; Rajib Podder, PhD; Albert Vandenberg, PhD; Diane M. DellaValle, PhD, RDN, LDN

Many thanks to the Reviewers for their time and comments. We have found all of these comments useful. As suggested, we have revised the manuscript, and addressed each issue here, point-by-point. We have submitted both MS Word Track changes and a Clean version for your convenience.
Reviewer #1:

1. There are many typos and grammatical errors - the paper needs a thorough proof-read.

   - We are sorry about the errors. We have carefully revised the manuscript.

2. In line 4 it was mentioned that "17.1% children (12-14 years)”; please clarify the age definition of children and source used.

   - We revised the line as “……….17.1% of adolescents aged 12-14 years were anaemic”.

   - References provided in [1].

3. Please use a consistent terminology for adolescent /children

   - We changed the word ‘children’ to ‘adolescent’ as appropriate.

4. Rational for conducting this study was not properly established. Background needs thorough revision.

   - This section has been revised.

5. Does the study aim to test any hypothesis? How much improvement in primary objective is expected?
- The study hypothesized that consumption of cooked iron-fortified lentil will increase body iron status and improve cognitive performance among the adolescent girls in Bangladesh. We stated the trial assumption in the ‘Trial Objectives’ section.

- We stated the expectation of the study in the first line of second paragraph under discussion section “Compared to the effect of iron supplementation to improve Fe status, the study expects a smaller effect size.”

6. Does this study include both married and unmarried adolescent girls??Please clarify

- The study included only unmarried adolescent girls. If an adolescent girl gets married after the baseline, she will be excluded from the study. This is because, we have observed from our personal experience that adolescents get pregnant after marriage and sometimes they are not aware about their pregnancy until after 2-3 months. In addition, the study was not designed to measure participants’ pregnancies.

- We added this in the manuscript as ‘…..non-smoking, not married, non-pregnant, not breastfeeding, and generally healthy at the time of enrollment.’

7. As this study will be conducted in "BRAC Adolescent Clubs" and the members include both boys and girls. In line 54 it was mentioned that the intervention will be offered to both boys and girls and the outcome will be assessed only among girls. Would you please explain why boys will not be included in the analysis??

- We did not include boys for the study because ‘Adolescents (particularly girls) are nutritionally vulnerable because of their significant growth, lifestyle and food habits, and more susceptible to Fe deficiency without anaemia due to menstrual losses of Fe [3,4].’ This statement was made in the earlier submitted version- last statement of the first paragraph under ‘Background’ section.

8. Inconsistency in trial registration information where it was mentioned that the trial was retrospectively registered while in table 1 suggest that the trial is not yet recruiting.
- We are sorry for the mistake. We deleted the word ‘Retrospectively registered’

9. What is the baseline prevalence of serum ferritin considered in the desired community and please mention reference??

- Community specific serum ferritin data is not available. We stated the statistics from Bangladesh National Micronutrients Status Survey (NMNS) report.

- This report is country representative and results are generalized. Study area was selected because it was one the study site used in Bangladesh National Micronutrients Status Survey (NMNS) 2011-2012.

- It was written as ‘A much lower prevalence of iron (Fe) deficiency (serum ferritin level <15.0 µg/L, 3.9% and 9.5%) was observed among children of age 6-11 years and adolescent of age 12-14 years, respectively, in the Bangladesh National Micronutrients Status Survey (NMNS) report [2].’ First paragraph under ‘Background’ section.

10. Randomization process does not clearly explain availability of total clusters in the sampling frame and how the study clusters will be randomized into three arms

- We revised the first part of ‘Sample size’ as follows “Considering the lower estimation of the expected difference in mean serum ferritin (5µg/L) with 80% power at p<0.05 significance level and inter-cluster correlation (ICC) at 0.025, a total of 16 clusters will be selected. Each cluster will have 3 clubs resulting in a total of 48 clubs. Within each club, n=27 eligible adolescent girls will be selected.”

- We added the following text under the “Sample size, randomization and blinding” sub-section under method section “Clubs will be randomly assigned to the intervention within each cluster. We did randomization at two level. Firstly, a total of 48 clubs will be randomly selected and clubs will be randomly assigned under 16 clusters. There will be 3 clubs in each cluster and 3
study arms (either iron fortified lentil or non iron-fortified lentils or no lentils) will be then randomly assigned within clusters using computer generated random assignments. Equal number of clubs (N 16) will be assigned to each arm and equal number of participants (N 420) will fall in each arm. Clubs, clusters and participants will be identified by sequential numbers such as 01, 02……48; 1001, 2002….1616 and 01, 02….1260 respectively.”

11. How the researchers will prevent spill of between intervention and control arms??

- We think, spill-over effect is less likely to occur as the study does not provide knowledge or any other information that may affect the study outcome. Study outcomes are biological and is to measure the changes in serum ferritin level and cognitive assessment among the adolescent girls.

- We added the following text in the discussion part “Spill-over effect is less likely to occur as the study does not provide knowledge or any other information that may affect the study outcome. Study outcomes are biological and is to measure the changes in serum ferritin level and cognitive assessment among the adolescent girls after consuming cooked amount of raw 37.5 gm lentils for 85 feeding days.”

12. Tablet mebendazole will be given in all arms or in intervention arms?? Please clarify.

- Tablet mebendazole will be given to all intervention arms just after the blood collection to ensure that adolescent girls be dewormed at the time the feeding starts. Before providing the antihelminthics, we will have their history of taking the medicine as it is not recommended to take deworming drugs within 4 months of previous consumption. We, therefore, plan to ask each adolescent girl on their deworming history (verified by their parents) and only if they are eligible, we will then provide to them.

- We replaced Tab Mebendazole (100mg) to Tab Albendazole (400mg). It is because Tab Mebendazole (100mg) chewable form is not available in Bangladesh. We therefore, replaced Mebendazole to Albendazole [Tab. Alphin® (400mg)] chewable form once to only those
participants who have given blood samples and no history to taking antihelminthics in last 4 months (verified by their parents).

- We added the following text in the discussion part “Tablet. Albendazole (400mg) will be provided so that participants are dewormed at the time of feeding trial meaning that they are free from parasites, such as roundworm, flukes and tapeworm infestation. Earlier studies suggested that worm infestation is linked with anaemia and deworming improves anaemic status among children [32-34].

13. Intervention description is not consistent throughout. Different components like rice with lentil, tablet mebendazole etc. are mentioned in non-coherent fashion.

- We added the following text in the discussion part, “Cooked iron-fortified lentil (Dal) is the only intervention considered in this study. To measure it’s efficacy, we had control other covariates that may influence the study outcome. To control other factors, we are providing non iron-fortified lentils, cooked rice is provided to increase consumption (based on the previous pilot study findings) and participants are dewormed by Tablet. Albendazole (400mg) at the time of feeding trial [32-34].

14. Exclusion criteria needs to include adolescents with severe anaemia.

- We will exclude those who are not generally healthy.

- As suggested, we will further exclude those participants who are severely anaemic. But it’d take around 1+ month to collect blood sample and analyze to detect those who are severely anaemic (hemoglobin <8 g/dl).

15. Overall, the write up requires a thorough revision to ensure consistency, flow and coherence.
- We are sorry the inconsistencies in the manuscript. We have carefully revised the manuscript.

Reviewer #2: This manuscript describes the protocol for a cluster randomised controlled trial looking at providing iron-fortified lentils to adolescent (aged 10-17) girls (n=1260) in Bangladesh. Outcomes include iron status (all participants, primary outcome), anthropometry (all participants), and cognitive performance (360 participants, primary outcome). The study is well motivated and mostly well described, although there are some important omissions around the randomisation procedure, details of blinding, what would count as evidence in favour of the fortified lentils, some details in the sample size information, and the statistical analyses.

- Thank you for your time to make comments. These all are useful and certainly enrich the quality of the paper.

Page 2, Line 4: Technically the study protocol describes a study with these aims.

- Yes, the study expects to measure 3 aims- iron status, anthropometry and cognitive performance.

Page 2, Line 4: Are you actually concerned with efficacy here (how well can it work under ideal conditions) or effectiveness (how well does it work under real-world conditions)? If the former, a per-protocol focused analysis would be standard with exclusion of non-compliers, if the latter, an intention to treat focused analysis would be standard with compliance not incorporated into that.

- This has been changed to Effectiveness throughout and analysis section is revised.

Page 13, Line 22 makes it clear there that it is effectiveness you are interested in, and not efficacy.

- Yes, there was disagreement within the Group about the use of the Terms – I do agree this is an Effectiveness trial and all has been changed throughout to reflect this.
Page 2, Line 6: "improving" compared to what, ordinary lentils or no lentils?

- We completed the sentence as ‘………improving body Fe status and cognitive performance compared to ordinary lentils and no lentils (usual intake) of non-pregnant adolescent girls in rural Bangladesh.’

Page 2, Line 10: The phrase "double blind" can be interpreted in many ways (see http://dx.doi.org/10.1136/ebm.7.1.4) and I think it would be useful to add the groups who will be blinded in parentheses here. Note that SPIRIT item 17a refer to exactly who will be blinded.

- Thank you for sharing the link. It was very useful.

- We revised the sentence as “….. double-blind (both trial participants and outcome assessors)……”

Page 2, Line 16: It’s not entirely clear what the primary question of interest here is. The description of the no-lentil group as the control would suggest that the key question is around the provision of lentils, but the iron-fortified group suggests to me that this is the key intervention here. Somewhere in the methods in the abstract, the reader should be able to find out what the primary research question is (are lentils useful for iron and cognition versus does fortifying lentils improve iron and cognition?) Line 40 below suggests that it might be the second of the questions that you are most interested in and this matches the online registration.

- We revised the second line of the background in the abstract ‘………improving body Fe status and cognitive performance compared to ordinary lentils and no lentils of non-pregnant adolescent girls in rural Bangladesh.

Page 2, Lines 19-22: The description of "a locally-acceptable recipe" suggests that monotony might be an issue. Could you add some information to the body of the manuscript covering this aspect?
Thank you for raising the issue. That was our concern too.

(1) In our earlier version, we stated the followings with citation under the ‘intervention’ sub-section under ‘Materials and Method’ section – “A locally-acceptable, standard lentil dal recipe identified during the earlier feasibility study will be used. For the thick cooking preparation of 100 g uncooked lentils, the recipe will include turmeric 5 g, chopped onion 40 g, garlic 8 g, green chili 3 g, water 700 ml, salt 1.5 teaspoon, soybean oil 10 teaspoons, and one small bay leaf (tejpata). The average cooking time would be around 18 min, and the approximate weight (after cooking of 37.5 g raw lentil) would be approximately 200 g [16].”

(2) We added the following sentences in the discussion part

“Other aspect is that as the study propose to serve same recipe for around 4 months, it is possible that single recipe for this length of time could lead to monotony among participants and may reduce the quantity of consumption or may face drop-outs after several days. The study predicted similar situation and addressed this in its pilot study which ran for 3 months (100 adolescent girls, 12 weeks, 5 days a week) and tested two different recipes (thick and thin preparation) in 3 different amounts (25 gm, 37.5 gm and 50 gm) using the local ingredients and cooking procedures in 2016-2017. This study came up with a single recipe as it is more acceptable by the similar participants (BRAC Adolescent Clubs in nearby ‘Gazipur’ district) and the pilot study did not find drop-outs or reduced consumption due to taste. Furthermore, it is important for the study to maintain its rigor and standardize the intervention to capture the true attribution of the feeding iron fortified lentils to the outcome compare to normal lentils and/or no lentils (usual intake). So, the probability of any fluctuation of the feeding trial due to the taste would affect equally to each intervention arm provided that we provide the same recipe for same duration to same participants. If we provide various recipe to increase the consumption, it is possible that either the study participants increase or decrease consumption compare to earlier recipe (even though probability of fluctuation remains the same) but that may influence the variation in study outcome.”

Page 2, Line 29: While it seems likely that the follow-up will be at 4 mo, this should be explicit.

- We revised lines as follows “Socio-demographic characteristics, household food security status, adolescent food habits and cognitive testing will be collected at baseline and endline at 4 months.”
Furthermore, we revised lines under the ‘Data collection tools and technique’ subsection (page 9) as follows- “Round 1 (baseline) will include all forms of data. Round 2 (midline) will only include blood sample data at 2 months. Round 3 (endline) will be the same as baseline at 4 months (Fig 1).”

Page 2, Line 31: While it seems likely that the midpoint interim measurement will be at 2 mo, this should be explicit.

- We added the following text “Venous blood samples will be collected at baseline (t=0 month), midpoint (t=2 months) and endline (t=4 months) to measure adolescents’ Fe status.”

Page 2, Lines 30-35: Is the cognitive testing at the interim or only baseline and follow-up times? (This is later clarified on Page 9, Line 12, but should also be clear here.)

- We revised lines as follows “Socio-demographic characteristics, household food security status, adolescent food habits and cognitive testing will be collected at baseline (t=0 month) and endline (t=4 months).”

- Furthermore, we revised lines under the ‘Data collection tools and technique’ subsection (page 9) as follows- “Round 1 (baseline) will include all forms of data. Round 2 (midline) will only include blood sample data at 2 months. Round 3 (endline) will be the same as baseline at 4 months (Fig 1).”

Page 2, Line 34: "carried-out"

- Corrected as suggested ‘carried-out’

Page 3, Lines 2-4: I suggest adding "of" after each of these percentages.
- We added ‘of’ after each of the percentages.

Page 3, Line 8: Suggest "ages" (add "s") here.

- Added ‘ages’. Thank you for point this out. We revised the manuscript carefully by an native English academician and corrected accordingly.

Page 4, Lines 10-12: Compared to ordinary lentils or compared to no lentils?

- We revised the sentence as “The trial assumes that the supplemental food-based Fe from the cooked Fe-fortified lentils will improve body Fe status and thus, cognitive performance compare to cooked ordinary lentils and no lentils.”

Page 4, Line 14: Same comment as above regarding efficacy/effectiveness.

- This has been changed to Effectiveness throughout.

Page 4, Line 16: Again, improving compared to what?

- We revised the lines as suggested “The primary objective of the study is to determine the efficacy of the Fe-fortified lentil-based dietary intervention compare to cooked ordinary lentils and no lentils in improving the Fe status of non-pregnant adolescent females in Bangladesh.”

Page 4, Line 32: Same comment as above regarding efficacy/effectiveness.

This has been changed to Effectiveness throughout.
Page 4, Line 38: Same comment as above regarding efficacy/effectiveness.

This has been changed to Effectiveness throughout.

Page 4, Line 41: Suggest "the drop-out" (adding "the"). Also, while the drop-out rate was 0%, were there missed meals during the study? This should be mentioned here.

- We added ‘the’ before the word ‘drop-out’.

- We revised the sentence as follows “The study findings suggested that adolescents were keen to eat cooked lentils over the 3 months intervention period and drop-out rate was 0%; although there were few missing meals.”

Page 4, Line 50: How often were the uneaten amounts below the 37.5g dry weight proposed here? Perhaps it would be best to give the amounts consumed for each group as means or medians. Also, just checking, this means that those in the 50g group consumed 12.5g more than those in the 37.5g group (if the uneaten amounts were the same)?

- We revised as follows “Higher palatability was observed for the thick preparation of cooked lentils compared to the thin preparation for all three intervention raw amounts of 25 g vs 37.5 g vs 50 g (cooked amount approx. 157.9 g vs 202.7 g vs 256.6 g respectively). There was no significant difference between uneated amounts of the served cooked lentil dish containing 37.5 g vs 50 g raw lentils (cooked amount approx. 202.7 g vs 256.6 g respectively.)”

Page 5, Line 2: Same comment as above regarding efficacy/effectiveness.

- This has been changed to Effectiveness throughout.

Page 5, Line 7: Same comment as above regarding blinding.
We revised the sentence as “….. double-blind (both trial participants and outcome assessors)……”

Page 5, Line 8: Same comment as above regarding efficacy/effectiveness.

- This has been changed to Effectiveness throughout.

Page 5, Line 8: Suggest "consuming" rather than "consumed"

- We changed the word ‘consumed’ to ‘consuming’.

Page 5, Line 14: Cluster RCTs don't eliminate selection biases. These still exist in the selection of the clusters and can arise in the consenting/assenting of participants within clusters (e.g. Page 5, Line 50). The main arguments in favour of cluster randomisation are convenience and minimizing contamination, with the price paid for these being reflected in the design effect and consequently larger sample sizes.

- We deleted the term ‘eliminates’ and replaced ‘minimize’.

- We revised the sentence as follows “Cluster-randomized controlled trials are the strongest design for making causal claims, as this design minimize bias from selection and confounding variables, allowing establishment of the temporal relationship. This design will further ensure better representativeness and prevent contamination of intervention between groups.”

Page 5, Line 14: Reliability is not a property of RCTs versus other designs. Reliability refers to whether the measured result would be the same if the measurement was repeated (at the individual or study level). For this at the study level, the sample size is crucial as is the reliability of the measurements at the individual level.
- We deleted the word ‘reliable’ from the above sentence.

Page 5, Line 22: Same comment as above regarding efficacy/effectiveness.

- This has been changed to Effectiveness throughout.

Page 5, Lines 22-26: I don't think that "BRAC—previously known as 'Bangladesh Rural Advancement Committee' is the #1 non-government organization (NGO) in the world for the third year running [17]." is relevant to the protocol and suggest its deletion. Adding a reference to the previous sentence would be appropriate though.

- We deleted the above sentence and revised as follows “The efficacy study will be conducted at the BRAC Adolescent Clubs (AC) [17].”

Page 5, Line 31: I suggest replacing "between ages" with "aged".

- We replaced "between ages" with "aged" as follows “Each club has a membership of 25-40 adolescent boys and girls aged 10-19 years.”

Page 5, Lines 46-50: This isn't a complete sentence, perhaps delete "who"?

- We revised the line “The study has been carefully designed to ensure that all of the adolescent girls between the ages of 10-17 years meeting the study’s inclusion criteria: non-smoking, not married, non-pregnant, not breastfeeding, and generally healthy.”

Page 5, Line 52: Given the clubs are for those aged 10-19 (Line 33 above), presumably those aged 18 or 19 will also be excluded from analyses?
- We moved the sentence to the later part of ‘Intervention’ subsection as follows “All adolescent boys and girls who attend the BRAC adolescent clubs will be offered the cooked lentils during the period of the intervention; however, the information of those girls not meeting the inclusion criteria will not be analysed.”

Page 7, Line 54: I think it is a little confusing to say "three lentil-based…arms" when one does not receive lentils.

- We are sorry for the mistake. We revised the line “There will be three arms in this efficacy trial.”

Page 7, Line 54: Same comment as above regarding efficacy/effectiveness.

- This has been changed to Effectiveness throughout.

Page 7, Line 58: Note again that the no-lentils group would be a control for the lentil-related questions but not the fortification ones. Note also that the point in the online registration that they can still consume lentils is important and should be incorporated here.

- This has been edited.

Page 8, Line 5: "lentils" (add "s")

- We added ‘s’ in the word ‘lentil’

Page 8, Line 28: Presumably this means "including an additional 20% to account for loss to follow-up”?
We revised the line as follows “Finally, n=420 adolescent girls will be included in each intervention arm including an additional 20% to account for loss to follow-up”

We revised the first part of ‘Sample size’ as follows “Considering the lower estimation of the expected difference in mean serum ferritin (5µg/L) with 80% power at p<0.05 significance level and inter-cluster correlation (ICC) at 0.025, a total of 16 clusters will be selected. Each cluster will have 3 clubs resulting in a total of 48 clubs. Within each club, n=27 eligible adolescent girls will be selected.”

Page 8: Lines 20-40: In order for the calculations here to be replicable, you also need to provide the SD of ferritin, the detectable cognitive effect size, the SD of this effect, and the number of clusters/number of girls per cluster for the cognitive sub-study. With this information, I should be able to produce the same numbers as you have. The number of girls needing to be invited should also incorporate a non-consenting/assenting rate(s). I am assuming that the study is powered for all three pairwise-comparisons? This would raise the issue of multiplicity. It is important that the "success criteria" for the study are made absolutely clear in the manuscript, either in this section or elsewhere. If the study affects Fe status only, will it be declared a success? What if only cognition shows evidence of response? Given that both these are listed as primary outcomes, my expectation is that unless both show evidence of improvement, the study will not be regarded as showing evidence in favour of iron-fortified lentils? This point needs to be made very clear in the protocol and will have implications for adjustments for multiplicity and so sample size.

Page 8, Lines 42-46: Same comment as above regarding blinding.

Page 8, Line 51: If I'm understanding correctly, the fortified and non-fortified lentils will be in different coloured bags but always of the same colour? If so, anyone reading this text would be
unblinded to the intervention (and so the colours must be deleted from the text). Using the same colours at all sites also means that if someone is unblinded at one site, all sites are potentially unblinded. If the colours vary between sites (yellow=fortified for some and non-fortified for others) or over time, that would strengthen this aspect of the design and should be made clear here.

- We revised it.

Page 9, Line 2: I'd suggest not capitalising "URGENT" here.

- We revised the format of the word as ‘urgent’

Page 9, Line 58-Page 10, Line 2: Unless this is going to be used to define a criterion for being included in analyses (must consumed a mean of 30g or more, for example), this is not an efficacy trial but is instead a pragmatic effectiveness trial. I think that effectiveness is the better question, but it does mean that the numerous references to efficacy would need to be changed if this is the case (which it does seem to be based on the statistical methods).

- This has been revised.

Page 10, Line 19: "detailed" (add "ed")

- We added ‘ed’ to the word ‘detail’.

Page 11, Lines 17-20: Presumably this intends to say that cognitive performance data will be collected from a subsample of the girls who provide venous blood?

- We revised the line as it was confusing “Secondly, survey data and cognitive performance data will be collected from those who will be given the venous blood sample.”
Page 11, Line 26: Are not the clubs the clusters? If these are different, this wasn't clear to me earlier.

- We added the following line in the “Sample size, randomization and blinding” sub-section
  “Considering the lower estimation of the expected difference in mean serum ferritin (5µg/L) with
  80% power at p<0.05 significance level and inter-cluster correlation (ICC) at 0.025, a total of 16
  clusters will be selected. Each cluster will have 3 clubs resulting in a total of 48 clubs. Within
  each club, n=27 eligible adolescent girls will be selected. Clubs will be randomly assigned to the
  intervention within each cluster. We did randomization at two level. Firstly, a total of 48 clubs
  will be randomly selected and clubs will be randomly assigned under 16 clusters. There will be 3
  clubs in each cluster and 3 study arms (either iron fortified lentil or non iron-fortified lentils or
  no lentils) will be then randomly assigned within clusters using computer generated random
  assignments. Equal number of clubs (N 16) will be assigned to each arm and equal number of
  participants (N 420) will fall in each arm. Clubs, clusters and participants will be identified by
  sequential numbers such as 01, 02……48; 1001, 2002….1616 and 01, 02….1260 respectively.”

Page 12, Figure 2: Why are cognitive assessments repeated at week 29? The figure also does not
show the interim assessments.

- We used the given SPIRIT template. Cognitive assessments will not be repeated at week 29.
  Fig 2 pointed study close-out at the week 29.

Page 13, Lines 7-10: More information is needed around anthropometry. What equipment is
being used (stadiometers, scales, tapes), how many measurements will be used, how will these be
combined, what is the measurement point for waist circumference? Ideally, quality control
should also be discussed here (e.g., inter- and intra-rater reliability).

- We added the following sentences as suggested “Participants’ height, weight, waist, hip and
  mid-upper arm circumference (MUAC) of adolescents will be measured to determine change in
growth [36-38]. Participants’ will be kept bare-footed, minimal clothing and we will be avoiding
carpets, sloping, rough, and uneven surface before the anthropometrics measurements. For
height, Frankfurt horizontal plane will be ensured, and they will be requested to put their heels

together. In addition, their backward curved body parts (buttocks and shoulder blades) and head

will be placed against the plane. For weight, participants’ will be requested to stand still putting

face forward and placing palm on their respective side and digital body weight bathroom scale

will be used after removing any sort of shoes and socks. Participants’ waist circumference will

be measured under the midline of armpit and at the midpoint between the inferior part of the last

rib and the top tip of the hip bone by using a constant tension tape. Hip circumferences will

be measured at the point of maximum diameter of buttocks using the same constant tension tape.

Same tape will be used to measure the MUAC. First, participants’ will be requested to put their

left arm at 90° angles, and midpoint will be marked between the distance of proximal and distal

point. Tape will be then wrapped around the point and measured ensuring that tape will be not be

either too tight or too loose. Measurement unit of all these anthropometric will be in centimeter

(cm) except weight will be captured in kilogram (kg). Participants’ mean Body Mass Index

(BMI) will be calculated by using BMI Percentile Calculator for Child and Teen [39].”

- We added the following sentences in the discussion part “Quality control will be ensured by

following multiple steps. For instance, direct training of enumerators, data quality control

supervisors, and data manager (all together by same trainers for same duration) will be carried

out prior to the commencement of data collection to adhere them to the consent, assent, and

questionnaire equally with the same degree of questioning and understand the format accurately.

Furthermore, in this face-to-face interview approach with using close-ended questionnaire,

response options will be formatted strictly. Field enumerators will be advised to cross-check the

questionnaire within team members before sending it to server using cellular internet data

connection. Additional training will be provided to the data quality control supervisors on field

level spot check of interview process and uncover any misunderstandings in the data collection

procedures. Each interview and arthrometric measurement will be taken by experienced female

interviewers separately in a private setting with presence of a witness. Each arthrometric

measurement will be taken 3 times to ensure consistency of the measurement. Survey and

dailywise data will be collected electronically by Open Data Kit (ODK) app on Android based

platform [40, 41]. This customizable mobile or tablet-based app could work on both online and

offline allowing to use GPS tracking, setting condition of the responses and enable real-time data

monitoring. An experienced data manager will be checking the received data thoroughly every

day and will be providing feedback directly to the enumerators. Additionally, data manager will

be assessing if there will be a need of re-training the field staff.”

Page 13: The randomisation method has not yet been described and must be added. See the

CONSORT guidelines for examples of the necessary information. The SPIRIT items 16a, 16b,

and 16c indicates pages 10 and 11 for these details, but I cannot see them there or elsewhere.
- We added the following text under the “Sample size, randomization and blinding” sub-section under method section “We did randomization at two level. Firstly, a total of 48 clubs will be randomly selected and clubs will be randomly assigned under 16 clusters. There will be 3 clubs in each cluster and 3 study arms (either iron fortified lentil or non iron-fortified lentils or no lentils) will be then randomly assigned within clusters using computer generated random assignments. Equal number of clubs (N 16) will be assigned to each arm and equal number of participants (N 420) will fall in each arm. Clubs, clusters and participants will be identified by sequential numbers such as 01, 02……48; 1001, 2002….1616 and 01, 02….1260 respectively.”

Page 13, Line 22: The clarification here that intention to treat will be used appears to confirm that this is an effectiveness and not an efficacy study.

- This has been changed to Effectiveness throughout.

Page 13, Line 28: You cannot use ANCOVAs or other simple statistical models here that assume independence between observations. Because of the clustering within clubs, you must use other options such as mixed models or generalised estimating equations. Given the interim measurements, these will also need to be incorporated into the analyses. This might be what you are referring to on Line 35 ("multilevel linear regression") but this needs to be clearer (what would be the random and what would be the fixed effects? Is there a residual covariance structure you have in mind for the longitudinal data?)
Page 13, Line 31: "as appropriate" is a very unclear phrase. What would make their inclusion appropriate?

Page 13, Line 33: Again, Chi-squared tests and logistic regression models assume independence and some form of modelling that incorporates the clustering within clubs is needed.

Page 13, Line 33: One of the main strengths of randomisation is that it eliminates confounding. While this can be introduced through differential attrition mechanisms, this needs to be addressed by a formal approach to missing data such as multiple imputation with perhaps selection or pattern-mixture models for informative missingness.

Page 13, Line 37: A brief description of model diagnostics that will be used should be added around here.

Page 13, Line 37: Same comment on confounding as above.
Page 13, Lines 36-39: A complete case analysis is not best practice. I'd strongly suggest that you look at MI and pattern-mixture or selection models to deal with missing data.

- This section has been revised.

Page 13, Line 44: It is traditional to also note the level of significance in this section.

- We added the following sentence “p < 0.05 will be considered statistically significant.”

Page 13, Line 50: Suggest deleting "the" before "adolescent girls".

- We deleted ‘the’ before the ‘adolescent girls’

Page 14, Line 36: These days, I think "layperson's" would be more appropriate.

- We replaced the term “layman’s” to “layperson’s”.

Reviewer #3: An interesting study protocol aiming to study the efficacy of iron fortified food supplementation in improving the iron status and cognitive performance amongst adolescent girls. The protocol is brilliantly designed and written. It methodically describes all aspects, nonetheless, following guidelines and adding the checklist is a Trials requirement. My observations to the manuscript are as below;

- Thank you for the comments.
Please explain how the study will ensure all aspects of data quality—reliability, correctness, completeness, timeliness, precision and integrity.

- We added the following sentences in the discussion part “Quality control will be ensured by following multiple steps. For instance, direct training of enumerators, data quality control supervisors, and data manager (all together by same trainers for same duration) will be carried out prior to the commencement of data collection to adhere them to the consent, assent, and questionnaire equally with the same degree of questioning and understand the format accurately. Furthermore, in this face-to-face interview approach with using close-ended questionnaire, response options will be formatted strictly. Field enumerators will be advised to cross-check the questionnaire within team members before sending it to server using cellular internet data connection. Additional training will be provided to the data quality control supervisors on field level spot check of interview process and uncover any misunderstandings in the data collection procedures. Each interview and arthrometric measurement will be taken by experienced female interviewers separately in a private setting with presence of a witness. Each arthrometric measurement will be taken 3 times to ensure consistency of the measurement. Survey and dailywise data will be collected electronically by Open Data Kit (ODK) app on Android based platform [40,41]. This customizable mobile or tablet-based app could work on both online and offline allowing to use GPS tracking, setting condition of the responses and enable real-time data monitoring. An experienced data manager will be checking the received data thoroughly every day and will be providing feedback directly to the enumerators. Additionally, data manager will be assessing if there will be a need of re-training the field staff.”

Due the small sample that too chosen from a specific set of population (BRAC adolescent club member) sample is non representative to rural massed, hence generalization of study results might be a challenge,

- We agree with the reviewer. The study aimed to established biological causality among the study participants and sample size was calculated accordingly.

Minor edits are required to remove the repetitiveness in the manuscript - inclusion and exclusion criteria, methodology, study setting etc.

- We apologize for the errors and revised the entire manuscript carefully.