PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort profile: The Shanghai PreConception Cohort (SPCC) for the association of periconceptional parental key nutritional factors with health outcomes of children with congenital heart disease
AUTHORS	Wang, Dingmei; Zhang, Yi; Jiang, Yuang; Ye, Ying; Ji, Mi; Dou, Yalan; Chen, Xiaotian; Li, Mengru; Ma, Xiaojing; Sheng, Wei; Huang, Guoying; Yan, Weili

VERSION 1 – REVIEW

REVIEWER	Léa Maitre ISGlobal
REVIEW RETURNED	15-May-2019

GENERAL COMMENTS	This study presents a unique opportunity to evaluate nutritional and clinical preconception factors and birth defects. Some gaps were identified in the manuscript as detailed below. Abstract Spell out abbreviations Genetic factors are mentioned but it is not mentioned again in the rest of manuscript how this was measured. Not mentioned, the women recruited during pregnancy, it should specifiy 6,573 couples of parents and 15,203 pregnant women Strength and limitations No limitations mentioned Who is in the cohort? Explained more about the 15,203 pregnant women recruited, justify they addition in a preconception study. Findings to date Add description of the outcome, birth defects, and distribution in this population Compare the population characteristics of the women included before conception and the early pregnant women in Table 1, with formal statistical testing. Figure 2 Link better the text of the figure and the text page 10, use same wording Page 9 population characteristics This information is repeated in page 15 but without reference to Table 1 and different numbers are presented. Please align the results and maybe only describe population characteristics in Findings to date.

	Collection of blood samples Are the samples kept within six hours after collection at room temperature? In findings to date, the outcome should be described in the recruited population, prevalence etc... Throughout the manuscript English language should be reviewed for accuracy.

REVIEWER	Marcia Feldkamp University of Utah USA
REVIEW RETURNED	29-May-2019

GENERAL COMMENTS	This is a manuscript that describes the Shanghai PreConception Cohort that is designed to prospectively collect information and blood from parents prior to conception to investigate nutritional factors and the risk for congenital heart defects (CHDs). The authors are to be commended for their success in developing and implementing this prospective cohort. Including objective markers (blood levels of nutrients) will be very important and informative in the future. They have also designed this cohort study to include a nested case-control study which will be of great value. The authors provide some basic demographic information on the enrollees but the paragraph on the findings to date are based on a small number of those participating overall. The only information that is useful is the percent of women using folic acid supplements prior to pregnancy. There is no biomarker data presented. Overall the manuscript is not structured such that it is easy for the reviewer and the grammar and sentence structure requires serious editing by someone with English as their first language. The introduction states that "To investigate the association between parental periconceptional key nutritional factors .. with the development of CHD and to explore the cutoff biomarker levels..." none of which are included in this manuscript. The challenge with this manuscript is that there is not yet sufficient data to warrant publication.

REVIEWER	Melissa Young Emory University
REVIEW RETURNED	03-Jul-2019

GENERAL COMMENTS	Manuscript provides an overview on the Shanghai PreConception Cohort. Impressive design that will provide novel data on preconception nutrition and birth defects. Abstract: -Please list the specific vitamins that will be examined. Avoid use of general terms nutrition/vitamins/nutrients without clearly defining what study will measure. Unclear at times if referring to dietary intake or biomarkers. -Additional details would be helpful on timing of data collection during preconception/pregnancy. -Specify key primary and secondary outcomes in abstract (what are the social and health outcomes?) Manuscript -What % of couples in Shanghai attend the preconception clinical

	visit? How do the characteristics of couples attending the preconception visit differ from the larger population delivering in local hospitals? This is briefly mentioned in limitations at end of manuscript but needs further elaboration. -Recommend adding a diagram to help illustrate the study participant flow (screening, enrollment, those meeting inclusion criteria, refusals, loss to follow-up, etc). Data available for the two cohorts of women enrolled during preconception vs during pregnancy -Please provide further details on the analytic plans (ex. association analysis does not provide sufficient information) -Unclear on “green channel in antenatal care”- please clarify? - Nutrition consulting is listed as a method of retaining participants in study. Please expand on services provided and how this may impact study findings. -Specify the volume of blood collected from participants. -How do methods for RBC folate compare to microbiological assay recommended by CDC? -Abstract mentions genetic, biological, social and other environmental exposures. Would be helpful to more clearly outline the key exposure variables and methods for data collection in each of these categories. -Page 15 mentions following children to school age but Figure 2 ends at delivery. Recommend clearly stating all key primary and secondary outcomes of the study. How is school age/study end defined? Further details are needed on data to be collected and research questions for long-term outcomes. -“Our findings suggest that education of folic acid supplement knowledge is deeply needed.” Beyond consumption data, were questions asked on folic acid knowledge, counseling, and receipt of supplements? What is known about the supply/availability of supplements? -More details are needed on the sub-sample of 553 samples in which the biomarker data is provided. How do the characteristics vary from the larger cohort of women? Caution over interpretation until full analysis of baseline data. Consider waiting until all baseline data are available for reporting. Manuscript would be greatly strengthened by having one strong manuscript that provides an overview of study design, research questions, methods, clear exposures and primary and secondary outcome and baseline descriptive data for full cohort. - A flow diagram on sample size and study recruitment/loss to follow up to date is needed. Table 1 -Is it >= college? -Provide currency for income Table 2 -unclear on “iron protein” Were any biomarkers for iron status assessed serum ferritin or TfR? -Were any biomarkers for inflammation (CRP/AGP) examined? This will be important for interpreting preconception micronutrient status. Figure 2 -Additional details on scope of questions and specific biomarkers at each time point would be helpful to add Appendix -typo “eath” each * Minor editorial/grammar edits needed throughout

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Léa Maitre
Institution and Country: ISGlobal
Please state any competing interests or state 'None declared':

Please leave your comments for the authors below
This study presents a unique opportunity to evaluate nutritional and clinical preconception factors and birth defects. Some gaps were identified in the manuscript as detailed below.

Abstract
1- Spell out abbreviations
Reply: We have spelt out abbreviations when it occurred first time, including "SPCC" and "RBC folate". (page 2 Abstract)

2- Genetic factors are mentioned but it is not mentioned again in the rest of manuscript how this was measured.
Reply: In the cohort, the genomic DNA of recruited subjects were extracted and stored for future study. Currently there are no candidate genes or variants genotyped. We have now provided the procedure for DNA extraction and storage in this manuscript (page 12, line 16 to page 13 line 4).

3- Not mentioned, the women recruited during pregnancy, it should specify 6,573 couples of parents and 15,203 pregnant women
Reply: Thank you for your comment. The numbers have been updated since our first submission. We modified the expression of recruitment "The baseline population included 8045 pre-conceptional couples, 3054 single women, and 15 615 early-pregnant women, respectively." We provided this information in the manuscript abstract, text (FINDINGS TO DATE, page 19, line 1 to line 15) and Tables.

Strength and limitations
4- No limitations mentioned
Reply:
We misunderstood the format requirements for this in the last version. We have now listed two limitations in the last paragraph of the manuscript in this version following the subheading "Strength and Limitations" on the page 19, line 8- page 20, line 2.
"Two limitations of this cohort study should be considered. Firstly, there are approximately 200 000 pregnant women giving birth every year in Shanghai, and approximately 20 000 of them will take part in the free preconceptional care in Shanghai, where participants were recruited consecutively. Although response rate was high (over 95%), pre-conception participants were recruited from a population voluntarily present in Shanghai city with pre-conception physical examination sites, who may have a stronger willingness for a healthy pregnancy. This may induce selection bias. Secondly, in this study, biological samples (cord blood, placenta) of the newborns are not collected. We plan to give new informed consent to the family who are willing to participate in future studies, to collect biological samples not mentioned before. In addition, electrochemiluminescence assay was used to examine serum and RBC folate concentrations, which is different from microbiologic assay that is used widely. This will not bias the association analysis but comparison with international populations needs caution.."

Who is in the cohort?
5- Explained more about the 15,203 pregnant women recruited, justify they addition in a preconception study.

Reply: Thank you for your comment. The numbers have been updated since our first submission. We modified the expression of recruitment “The baseline population included 8045 pre-conceptional couples, 3054 single women, and 15 615 early-pregnant women, respectively.
The recruitment of pre-pregnancy subjects was slower than expected, and the proportion and duration of become-pregnant among which within a certain period of time were uncertain. Therefore, at the beginning of 2017, we decided to recruit a group of early-pregnant participants (below 14 gestational weeks) in order to investigate key nutrient levels that are close to the critical window of fetal organ development. Please see the descriptions on page 6, line 20 to page 7, line 5.

Findings to date
6- Add description of the outcome, birth defects, and distribution in this population
Reply: We have add three results in the “finding to date” section. (1) The size of the established baseline sample and achieved birth data. By the end of November 2018, the last subjects recruited for the early-pregnant sample were due for delivery, however, by now we have achieved birth records of 12402 newborns. The follow-up of outcomes of the rest of the participants is ongoing. The number of pregnancies among participants recruited before pregnancy is under investigation. (2) Achieved number of CHD outcome. A total of 151 cases of congenital heart disease were identified through the Shanghai CHD screening platform, of which 20 cases were from pre-pregnancy samples and 131 cases were from pregnancy samples. The prevalence of CHD is 10.5‰(131/12402) based on the present available data. (shown in Figure 2) (3) A summary description of some biomarkers examined in a small pilot study (page 15 line 10- page 17, line 10). According to our design, only selected subjects by nested case-control design according to specific outcomes will be examined for the full list of biomarkers.

7- Compare the population characteristics of the women included before conception and the early pregnant women in Table 1, with formal statistical testing.
Reply: Thank you for the comment. We made comparisons between the two groups of females showing that they were similar in age but different in education levels and occupation. The prevalence of smoking and alcohol drinking were much lower in early-pregnant women (Table 1). This information has been added to the text (page 15 line 10 to page 17, line 10). As we state, this information was partly included in our another manuscript which was recently submitted to Public Health Nutrition with Manuscript number of PHN-RES-2019-0914.

Figure 2
8- Link better the text of the figure and the text page 10, use same wording
Reply: Thank you for your comment. We modified descriptions of Figure 2 in the text to make it consistent to the figure legends.

Page 9 population characteristics
9- This information is repeated in page 15 but without reference to Table 1 and different numbers are presented. Please align the results and maybe only describe population characteristics in Findings to date.
Reply: Thank you for your comment. We moved the description of population characteristics to the “FINDINGS TO DATE” section.

Collection of blood samples
10- Are the samples kept within six hours after collection at room temperature?
Reply: In this study, the rest blood samples for routine clinical examination were collected. The blood sample for routine clinical examination was usually 5 ml and extracted in the morning. Routine clinical examination followed which was performed at room temperature. The rest blood samples (fasting serum and EDTA anticoagulation) of peripheral venous blood from routine laboratory examination were kept. These blood samples were temporarily stored in a 4℃ refrigerator for dispensing within 6

hours and transferred to a -20 degree freezer. After completion of blood sample distribution the serum and the whole blood were stored at the site laboratory and then transported by three trained investigators to the central biobank for storage in -80 ℃ freezers within two weeks. During the collection and transfer process samples were labeled and recorded in the sample system. In order to detect chemicals (folate) that are sensitive to light sampling tubes were made of a light-proof material and the process of collecting blood samples were completely protected from light. (page 10, line 4-19)

11- In findings to date, the outcome should be described in the recruited population, prevalence etc...
Reply: Thank you for your suggestion. We updated this section in this version based on the currently available outcomes.
Now, three main results were presented in the “FINDINGS TO DATE” section. (1) The size of established baseline sample and achieved birth data. By the end of December 2018, the last subjects recruited for the early-pregnant sample were due to delivery, however, by now we have achieved birth records of 12 402 newborns, the follow-up of outcomes of the rest of participants is ongoing. The number of pregnancies among participants recruited before pregnancy is under investigation. (2) Achieved number of CHD outcome. A total of 151 cases of congenital heart disease were identified through the city CHD screening platform, of which 20 cases were from pre-pregnancy samples and 131 cases were from pregnancy samples. The prevalence of CHD is 10.5‰(131/12402) based on the present available data. (shown in Figure 2) (3) A summary distribution of some biomarkers in a small pilot study (page 16, line 7 -14). According to our design, only selected subjects by nested case-control design according to specific outcomes will be examined for the full list of biomarkers.

12- Throughout the manuscript English language should be reviewed for accuracy.
Reply: Thank you for your suggestion. We corrected the grammar mistake, and asked a company for the help of a professional copyediting service.
Reply: We have asked for help from a native speaker to improve the language editing throughout the manuscript.
Reviewer: 2
Reviewer Name: Marcia Feldkamp
Institution and Country: University of Utah, USA
Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
-This is a manuscript that describes the Shanghai PreConception Cohort that is designed to prospectively collect information and blood from parents prior to conception to investigate nutritional factors and the risk for congenital heart defects (CHDs). The authors are to be commended for their success in developing and implementing this prospective cohort. Including objective markers (blood levels of nutrients) will be very important and informative in the future. They have also designed this cohort study to include a nested case-control study which will be of great value.
The authors provide some basic demographic information on the enrollees but the paragraph on the findings to date are based on a small number of those participating overall. The only information that is useful is the percent of women using folic acid supplements prior to pregnancy. There is no biomarker data presented.

Overall the manuscript is not structured such that it is easy for the reviewer and the grammar and sentence structure requires serious editing by someone with English as their first language. The introduction states that "To investigate the association between parental periconceptional key nutritional factors. With the development of CHD and to explore the cutoff biomarker levels..." none of which are included in this manuscript.
The challenge with this manuscript is that there is not yet sufficient data to warrant publication.

Reply: This cohort profile aims to introduce the cohort protocol and currently available baseline data.

We understand your interest in the biomarker data. We conducted a pilot study in a small subgroup of subjects, bio-samples were examined for some biomarkers. Summary descriptions were presented in Table 2 as a part of “FINDINGS TO DATE” section. The distribution of these biomarkers and their differences between CHD cases and matched controls will be the main results of our next manuscript “association of maternal and paternal peri-conceptional nutrition factor levels with offspring CHD”.

Table 2. Distributions of biomarkers in women before and at early gestation (pilot study)
Biomarker PreConception Early pregnancy
n level n level
Serum folate, ng/ mL 620 9.7(6.5, 13.8) 577 14.5(11.2, 16.4)
RBC folate, ng/ mL 570 247.0(184.8, 340.5) 587 417.4(308.6, 544.2)
Homocysteine, μmol/L 624 6.5(5.2, 8.6) 599 4.2(3.5, 5.2)
Vitamin B12, pg/mL 625 495.2(394.2, 639.0) 600 388.5(289.4, 511.4)
Vitamin D, ng/ mL 607 16.3±6.0 578 15.5±6.1

Reviewer: 3
Reviewer Name: Melissa Young
Institution and Country: Emory University
Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
Manuscript provides an overview on the Shanghai PreConception Cohort. Impressive design that will provide novel data on preconception nutrition and birth defects.
Abstract:
1- Please list the specific vitamins that will be examined. Avoid use of general terms nutrition/vitamins/nutrients without clearly defining what study will measure. Unclear at times if referring to dietary intake or biomarkers.
Reply: Thank you for your comment. Due to the word limit of the abstract, we are unable to specify all measured biomarkers or dietary intake in detail. However, we list the biomarkers (page 10-14,) and the content of the questionnaire (page 9-10,) in the main text.
Considering your suggestion, we modified the abstract that “Multiple nutrition factors in blood sample of participants that were selected by case-control design will be examined, including folates, homocysteine, vitamin A, vitamin D, vitamin E and metals.”

2- Additional details would be helpful on timing of data collection during preconception/pregnancy.
Reply: Thank you for your suggestion. The timings of collecting data of blood sample and questionnaire are showed in Figure 1.

3- Specify key primary and secondary outcomes in abstract (what are the social and health outcomes?)
Reply: Thanks. We rewrote the objectives of the abstract as follows:
The Shanghai PreConception Cohort (SPCC) was established initially to investigate associations of parental periconceptional nutritional factors with congenital heart disease (CHD), and has extended to children growth, development and pediatric diseases. (page 2 Abstract)

Manuscript
4- What % of couples in Shanghai attend the preconception clinical visit? How do the characteristics of couples attending the preconception visit differ from the larger population delivering in local hospitals? This is briefly mentioned in limitations at end of manuscript but needs further elaboration.
Reply: Your question is about the representativeness of the study population. Based on official historical data of year 2014, about 1/5 couples in Shanghai city (40 000/200 000 live birth) attend the

preconception clinical visit.
The recruitment of our study were conducted at the city pre-conception examination sites. This may induce selection bias that subjects were limited to the participants who volunteered attending the examination, and may have had a stronger willingness to have a health pregnancy. The nutritional supplement usage rates and the level of nutrition biomarker may be higher than Shanghai population. We added this into the “Strengths and Limitations” in the Discussion section on the page 19, line 1-15.

5- Recommend adding a diagram to help illustrate the study participant flow (screening, enrollment, those meeting inclusion criteria, refusals, loss to follow-up, etc). Data available for the two cohorts of women enrolled during preconception vs during pregnancy
Reply: Thank you for your suggestions. We added a flow chart to illustrate the study participant flow (Figure 2).

6- Please provide further details on the analytic plans (ex. association analysis does not provide sufficient information)
Reply: Thank you for your suggestions.
We added a brief statistical analysis plan on page 14, line 5 to page 16, line 9. The complete integrated statistical analysis plan will be uploaded to the clinical trial website. (www.clinicaltrials.gov)

“To investigate the association of maternal pre-conception nutrition levels with offspring CHD risk, a nested case-control study will be conducted. The control will be matched by age and site.
The sample size for the nested case-control analysis was planned as 180 cases and 720 matched controls to detect a maternal folate deficiency with prevalence of 50% in controls with odds ratio of 1.6 in association to achieve a power of 80% at alpha of 0.05. Based on CHD incidence above 8.94 per 1 000 live births[4], 20 000 pregnancies will be needed. For a continuous nutrient variable with standard deviation 2.0, 50 matched-pairs (1:4) are required to achieve 90% power to detect an odds ratio of 1.3 calculated using conditional logistic regression with a 0.05 significance level[23, 24]. Once a sufficient number of CHD cases is achieved, the quantitative association of pre-conceptional RBC folate levels with CHD using nested case-control design will be investigated.
Conditional multivariate logistic regression will be used for association analysis with offspring affected status of CHD being the dependent variable, nutrition factors levels as exposure and adjusted for all potential paternal and maternal covariates. Odds ratios (OR) and 95% confidence intervals (95%CI) will be reported. To explore a potential cutoff point of the nutrition levels that significantly increases the risk of CHD, a dummy variable will be set up by categorizing the maternal pre-conception nutrition levels based on the distribution of the control group. The does-response relationship will be also be analyzed. Sensitive analysis will include non-conditional logistic regression analysis, or generalized estimation equations (GEE) model, or generalized linear models when necessary.”

7- Unclear on “green channel in antenatal care”- please clarify?
Reply: It is quite common in China that patients need to line up for long time in the big hospital. The green channel means that we set specific procedures and arrangements to our subjects to see a doctor quickly. This will be attractive to participants.

8- Nutrition consulting is listed as a method of retaining participants in study. Please expand on services provided and how this may impact study findings.
Reply: We meant that we provided a phone number for the research team in the contact card given to the participants. This is not for providing nutrition consulting, but to answer the possible questions from participants. We changed the sentence to avoid misunderstanding as the following “We also provided a contact number on the participant card to answer their calls for queries about the study procedures..”(page 9, line 3-4) .

9- Specify the volume of blood collected from participants.

Reply: Thank you for your suggestion. The routine clinical examinations need 5 ml blood per tube. We kept the rest of fasting serum and EDTA anticoagulation blood samples of peripheral venous blood after routine laboratory examinations. We specify the volume that "…The blood sample for routine clinical examination is usually 5 ml…" in page 10, line 6.

10- How do methods for RBC folate compare to microbiological assay recommended by CDC?
Reply: Microbiological assay is a classic method for detecting folic acid in living organisms and is recommended by US CDC for testing serum folates. Electrochemiluminescence assay has been widely used in clinical practice. The roundtable summary also supported the use of the LC-MS/MS procedure for the measurement of serum folate in future NHANES but agreed that this procedure is not yet fully validated for the measurement of RBC folates. We discuss this as one of a limitation in our manuscript as following:
"Electrochemiluminescence assay was used to examine serum and RBC folate concentrations, which is different from microbiologic assay that was used widely. This will not bias the association analysis but comparison with international populations needs caution."
The reference:
Yetley EA et al. Biomarkers of folate status in NHANES: a roundtable summary. Am J Clin Nutr. 2011 Jul;94(1):303S-312S. doi: 10.3945/ajcn.111.013011.

11- Abstract mentions genetic, biological, social and other environmental exposures. Would be helpful to more clearly outline the key exposure variables and methods for data collection in each of these categories.
Reply: Thanks. Following your suggestion, we modified the abstract to "General characteristics, routine clinical data, consumption of diet supplements, such as folic acid and compound vitamins were collected, serum and plasma samples were collected at pre-conception, early, middle and late gestation respectively by standard procedures. Multiple nutrition factors in blood sample of subjects that were selected by case-control design will be examined, such as folates, homocysteine, vitamin A, vitamin D, vitamin E and metals. Genomic DNA were extracted.(page 2 Abstract)
We have presented detailed information about biomarkers (page 10: Examination of key nutrition factors in blood samples) and the questionnaires (page 9-10: Personal characteristics questionnaires) in the main text.

12- Page 15 mentions following children to school age but Figure 2 ends at delivery. Recommend clearly stating all key primary and secondary outcomes of the study. How is school age/study end defined? Further details are needed on data to be collected and research questions for long-term outcomes.
Reply: Thank you for your suggestion. Yes, we declare that the Shanghai PreConception Cohort (SPCC) is a long-term cohort and meets the "Cohort profile" requirement. We have a complete plan to follow-up offspring to the age of 18 years old. There will be three phases, peri-natal (pre-conception, conception and birth), 0-6 years (infants and pre-school age), and 7-18 years (school age). The cohort will be financially supported by different grants. The current manuscript focuses on the first phase, the establishment of the baseline and our first main outcome, the congenital heart disease (CHD). This information is presented in a new figure (Figure 1). The data collection plan for the 2nd and 3rd phase are included Appendix 4. Multiple outcomes including growth, physical and neuro-developments, obesity, and common and rare diseases will be investigated. The above descriptions can be seen in the text (on page 17: FUTURE PLANS).

13- "Our findings suggest that education of folic acid supplement knowledge is deeply needed." Beyond consumption data, were questions asked on folic acid knowledge, counseling, and receipt of supplements? What is known about the supply/availability of supplements?
Reply: We did not investigate the folic acid knowledge, counseling, and receipt of supplements, only consumption information is investigated by the questionnaires. Perinatal health care in China

recommends diet supplementation are provided to every couple when they register for marriage and preparing for pregnancy with printed handouts, which is commonly included as primary community-based health care providers. However, it is their right to choose their favorite supplements products. Folic acid supplements are commonly available in pharmacies in China. We changed the sentence into “Our findings suggest that a greater emphasis for use of folic acid supplement before pregnancy is deeply needed.”

14- More details are needed on the sub-sample of 553 samples in which the biomarker data is provided. How do the characteristics vary from the larger cohort of women? Caution over interpretation until full analysis of baseline data. Consider waiting until all baseline data are available for reporting. Manuscript would be greatly strengthened by having one strong manuscript that provides an overview of study design, research questions, methods, clear exposures and primary and secondary outcome and baseline descriptive data for full cohort.
Reply: Thank you for your comments. Currently based on available data, we have two main findings. Firstly, only 42% women take folic acid supplementation before pregnancy, based on questionnaire data. This result is included in another manuscript which is under review by Public Health Nutrition. Secondly, the RBC folate levels are low. Only 15% prepare-for-pregnant women and less than half early-pregnant women had RBC folate concentration meeting the recommended levels for preventing NCTs. This finding is based on a small subgroup of women (n=652 for pre-pregnant and n=557 for early-pregnant women, the numbers are updated to latest) consecutively enrolled in one single site during a short period of time. We added some basic information of the subgroup in the text (page 16 line 15- page 17 line 7) and rewrote the part as following:
“We conducted a small pilot study in April 2017 to explore blood levels of nutrition factors, including serum folate, RBC folate, vitamin A, vitamin E, and vitamin D. The blood samples from 627 females were selected consecutively from the preconception sample according to who was identified pregnant. In addition, 597 women who were consecutively recruited from the antenatal care clinics were selected. As shown in Table 2, the concentrations of RBC folate median level is 247.0 ng/ml (IQR: 184.8-340.5 ng/ml) in preconception women and 417.4 ng/ml (IQR: 308.6-544.2 ng/ml) in early-pregnant women. Twenty percent of preconceptional participants and 44.9% of pregnant participants had a folate level over 400 ng/ml, which was suggested as optimal level for preventing neural tube development defects [25, 26]. These results suggest that effort is urgently needed to improve the intake of folic acid supplementation in the prepare-for-pregnancy population, especially before pregnancy.” (page 16 line 15 to page 17 line 7)

15- A flow diagram on sample size and study recruitment/loss to follow up to date is needed.
Reply: Thank you for your suggestion. We agree with you and added a flow chart to represent the research process. (shown in Figure 2)

16- Table 1
Is it >= college?
Provide currency for income
Reply: Thanks. We corrected “>college” with “>=college”. The currency for income is RMB with unit of ￥. We add the signal “￥” in the Table 1.

Table 2
17- unclear on “iron protein” Were any biomarkers for iron status assessed serum ferritin or TfR?
Reply: Thanks. We are sorry for the typo. We measured “serum ferritin”. It is changed on page 11, line 5 Because of the content of Table 2 is similar with Figure 1, and we deleted the previous Table 2.

18- Were any biomarkers for inflammation (CRP/AGP) examined? This will be important for interpreting preconception micronutrient status.

Reply: Thank you for your suggestion. The crp/agp was not included in the routine examination in pre-pregnancy clinics. We will consider your suggestion. We are able to test inflammatory makers when necessary based on our stored serum sample.

Figure 2
19- Additional details on scope of questions and specific biomarkers at each time point would be helpful to add
Reply: Thank you for your comment. The frame of the design of SPCC is shown in the revised currently in Figure 1.
We added a Table 3 to describe in detail the possible of research questions and specific biomarkers based on collected type of bio-samples.

Table 3. Biosample collected and biomarkers that can be examined in the SPCC cohort
Biosample Available in subjects Available Sample type and volume Time
Preconception+early pregnancy
(Baseline) 24-28 weeks 32-36 weeks
Mother (n=25487) (n=8668) (n=7522)
Serum, 200 ul*3 Yes Yes Yes
Whole blood Yes Yes Yes
Genomic DNA, 150 ng Yes Yes Yes
Father (n= 7151) - -
Serum, 200 ul*3 Yes NA NA
Whole blood Yes NA NA
Genomic DNA, 150 ng Yes NA NA
Child NA
Scope of research questions:
1. Quantitative association of pre-conceptional key nutrition levels (such as serum and RBC folates and related biomarkers, such as homocysteine and vitamin B12) with CHD, currently, and other folate sensitive birth defects.
2. Quantitative association of peri-conceptional maternal and paternal key nutrition levels (preconception levels, and dynamic levels during gestation) with important maternal and neonatal gestational complications, neuro-development of infants, childhood obesity, and clinical pediatric diseases.
3. Periconcpetional maternal and paternal folate concentration with Autism spectrum disorder, allergy and asthma in children.
Biomarkers that will be examined in different type of bio-samples:
1. Biomarkers based on serum sample:
a) Folate and related markers: serum folate, homocysteine
b) Other Vitamin: vitamin A, vitamin B12, vitamin E, and vitamin D
c) Marco and micro metals: Mg, Fe, Zn, Se, Mn, As, Cu, Ca, ect), mg/L
d) Serum ferritin
e) Fasting glycemic and lipid profiles: fasting glucose, total cholesterol, triglyceride, high density lipoprotein, low density lipoprotein
2. Whole blood sample: RBC folate
3. Genomic DNA sample: candidate genetic variants or genome-wide variants are all possible to examined.

20- Appendix
typo “eath” each
* Minor editorial/grammar edits needed throughout

Reply: Thanks. We corrected it in the Appendix and have asked a company for professional copyediting service.

VERSION 2 – REVIEW

REVIEWER	Léa Maitre ISGlobal, Spain
REVIEW RETURNED	27-Aug-2019

GENERAL COMMENTS	Some English language spelling should be done in the new parts amended and the abstract. Table 1. specify the statistical test used under te table Table 2. specify what is the level (average and IQR?) Future plans Specify in the text the different time windows (6 weeks, 6 months etc,) the number of repeat assessments, and the type of assessments in further details, at least specifying the clinical domains of interest: neurodevelopment, cardio-vascular, etc... Figure 2 should state the numbers known by December 2018 for the follow up for the number of pregnancies and deliveries.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Léa Maitre
Institution and Country: ISGlobal, Spain
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
Q1: Some English language spelling should be done in the new parts amended and the abstract.
Answer 1: Thanks for your suggestions. We had carefully checked the manuscript, and then had asked for help from Dr. Asif Naseem for English editing again. We have made some spelling corrections in the manuscript.

Q2: Table 1. specify the statistical test used under the table
Answer 2: Thanks for your suggestions. We made the following corrections in the footnotes of Table 1:
t tests were used to compare numerical variable (age). Chi-square tests were used to compare categorical variables (ethnicity, educational level occupation, smoking and alcohol drinking).

Q3: Table 2. specify what is the level (average and IQR?)
Answer 3: Thanks for this useful comment. Median (IQR) was used to describe the levels of serum folate, RBC folate, homocysiteine and vitamin B12. Mean ±SD was used to present the level of vitamin D with normal distribution. We have added "Median (IQR)" and "Mean ±SD" for these variables in Table 2

Q4: Future plans
Specify in the text the different time windows (6 weeks, 6 months etc,) the number of repeat assessments, and the type of assessments in further details, at least specifying the clinical domains of interest: neurodevelopment, cardio-vascular, etc...
Answer 4: Thanks for your suggestions. During the stage of 0 to 6 years old, the data of neurodevelopment from routine child care clinical visit at birth, 6 weeks, 6 months, 12 months, 36 months, and 60 months through the cooperating medical institutions. Physical measurements data and dietary intake information also can be collected at this stage. During the stage of 6 to 18 years old, we plan to follow up their growth (height, weight, blood pressure, et al.) mainly relying on the annual physical examination of Shanghai Student Health and Fitness Surveillance Center System. Multiple outcomes for children, including growth and development, cardiovascular diseases,

neurodevelopment, metabolic diseases, obesity, and hypertension will be investigated. Please see Appendix 4 for details. (Page 17, line 20 to Page 18, line 10).

Q5: Figure 2 should state the numbers known by December 2018 for the follow up for the number of pregnancies and deliveries.
Answer 5: Thanks. We have achieved birth records of 12 402 newborns in pregnant sample by the end of December 2018, but not “by now”. We have modified the sentence to “By the end of December 2018, the last participants recruited at early pregnancy were due for delivery, however, we have achieved birth records of 12 402 newborns.” (Page 16, line 9 to 12). The data of pregnancies and deliveries in preconception sample is extracted from Maternal Clinic Antenatal Medical Record System, which is administrated by Shanghai municipal government. The latest data extraction was in May 2018, by which, the number of pregnancies and delivers were 1538 and 975 respectively. Next extraction is planed at the end of 2019. Following your suggestion, we added these numbers and stated the correct timing in Figure 2.