PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort profile: the Westlake BioBank for Chinese (WBBC) pilot project
AUTHORS	Zhu, Xiao-Wei; Liu, Ke-Qi; Wang, Ping-Yu; Liu, Jun-Quan; Chen, Jin-Yang; Xu, Xue-Jin; Xu, Jin-Jian; Qiu, Mo-Chang; Sun, Yi; Liu, Chu; Bai, Wei-Yang; Zhao, Pian-Pian; Xia, Jiangwei; Gai, Si-Rui; Guan, Peng-Lin; Qian, Yu; Cong, Pei-Kuan Cong; Xie, Shu-Yang; Zheng, Hou-Feng

VERSION 1 – REVIEW

REVIEWER	Rush, Amanda The University of Sydney Sydney Medical School
REVIEW RETURNED	15-Jan-2021

GENERAL COMMENTS	This paper describes early results from a cohort study of Chinese adolescents. Whilst comprehensive data on phenotype and lifestyle factors have been collected, the paper lacks clarity. The paper is not presented in a typical IMRAD style, and the writing would benefit from a proof read to improve the English. The message is vague and without novelty. In particular, it is not clear in the abstract and introduction what the scope of the results and discussion are for this particular paper. For example, there is mention of the genomics data collected on a subset of participants, however there are no genomics results. The Introduction does not lead the reader to what is to come in the results and disucssion, but rather reads as three sections from a text book. There are more than one instance where the text is overly vague eg. Page 8 - exclusion criteria include 'participants missed most of the items in data collection'. Additionally, the methods are vague in some instances eg. (page 9) - to ensure accurate data participants were asked to have no 'excessive physical activity for at least one night'. There is also no definition for smoking status (just yes or no?) or alcohol consumption (just current drinker or not?). Figure 1 is particularly difficult to read/interpret and doesn't add much value to the paper. There is no real discussion in the paper at all. I suggest waiting until longitudinal data provides some scientific novelty, before creating a narrative with clarity in the results, and a strong, separate discussion of the findings.

REVIEWER	Olson, Janet Mayo Clinic, Epidemiology
REVIEW RETURNED	22-Jan-2021

GENERAL COMMENTS	1) The WBBC study is a well-designed cohort study with strong likelihood of providing meaningful data in the future. The use of the word 'pilot' in the description of this cohort throughout this manuscript is rather confusing. A pilot study implies that you intend to expand this cohort to a larger # of subjects, but this was not discussed anywhere. I recommend dropping the word "pilot" and just refer to the WBBC study. If you are referring to the 3 recruitment phases, perhaps use "phase" rather than "pilot". 2) Providing #s to the 2nd decimal is meaningless. Round to 1 decimal. 3) Use males/females consistently rather than boys/girls, men/women. The latter implies an age stratification that I do not believe the authors intend. 4) Minor, but frequent, corrections to English grammar are needed, including referring to adolescence when it should be adolescents, missing plurals, missing articles (a, the). I have included a pdf of my edits/suggestions that are recommendations and not likely exhaustive. 5) Inclusion/exclusion criteria should be improved to be more clear. Was it required that all complete all the exams, BMD scans, blood tests and questionnaires? Please clarify. Also, there is no reason to include the inverse of an inclusion criterion as an exclusion criterion. 6) Line 323 – clarify which phenomenon to which you are referring 7) Lines 360-365 – this is a run-on sentence with multiple ideas. Divide into fewer sentences to clarify and simplify. 8) Table 5 – Improve the variable descriptions so it is clear which measures and cutpoints were used to assign subjects to each group

VERSION 1 – AUTHOR RESPONSE

Reviewer #1
Comments to the Author:

Dr. Amanda Rush, The University of Sydney Sydney Medical School

Comments to the Author:

This paper describes early results from a cohort study of Chinese adolescents. Whilst comprehensive data on phenotype and lifestyle factors have been collected, the paper lacks clarity.

Re: Thanks so much for your comments, please see our responses below.

The paper is not presented in a typical IMRAD style, and the writing would benefit from a proof read to improve the English.

Re: Thanks, this is a “Cohort profile” paper with the format provided by the journal. And we have read through the paper carefully and corrected some grammar errors.

The message is vague and without novelty. In particular, it is not clear in the abstract and introduction what the scope of the results and discussion are for this particular paper. For example, there is mention of the genomics data collected on a subset of participants, however there are no genomics results.

Re: Sorry that we didn’t make it clear. The Westlake BioBank for Chinese (WBBC) is an open-ended prospective study with very broad research aims, an extensive range of data collection in the WBBC study. This paper is a descriptive study of the pilot project of WBBC, the main purpose is to profile the cohort, most of the results are descriptive, therefore, in this paper, we demonstrated the sampling design of the participants, the data collection procedures (including the measurements of anthropometric parameters, biochemistry assessment and questionnaire-based assessment), and presented the descriptive results of the measurements, and tried to answer: (1) what is the prevalence of underweight, overweight, obesity and vitamin D deficiency in Chinese late adolescence? What is the reference value of serum vitamin D level with the young people? (2) what is the difference between male and female in term of height, weight, blood pressure, lifestyle and bone health in the young people.

And we have got progress on our sequencing data, the genomic study has been carried out and been described elsewhere (https://www.biorxiv.org/content/10.1101/2021.02.06.430086v1). In brief, we demonstrated the Chinese genetic structure for Chinese population and we provided an online imputation server (https://imputationserver.westlake.edu.cn/) which could result in higher imputation accuracy compared to the existing panels. We have reported our sequencing progress in the maintext (please see page 11, line 258 to line 269).

The Introduction does not lead the reader to what is to come in the results and discussion, but rather reads as three sections from a text book.

Re: Sorry about this, we have re-organized and re-written the ABSTRACT. In brief, we introduced what is Westlake BioBank for Chinese (WBBC) at the first pharagraph, and what is the pilot project of WBBC, we then proposed what specific questions this paper would like to answer. In general, WBBC is designed as a large-scale prospective cohort with its aim of recruiting at least 100,000 Chinese samples at different age. The pilot project of WBBC has focused on the study on the young population (Late adolescence), and has already collected a wide range of information including demographics and anthropometric measures, serological tests, physical activity, sleep quality, age at menarche and bone mineral density etc. The main purpose of this particular paper is to profile the cohort, therefore, only limited findings were reported and few questions were asked, e.g. (1) what is the prevalence of underweight, overweight, obesity and vitamin D deficiency in Chinese late adolescence? What is the reference value of serum vitamin D level with the young people? (2) what is the difference between male and female in term of height, weight, blood pressure, lifestyle and bone health in the young people. Please see page 2, line 2 to line 27.

There are more than one instance where the text is overly vague eg. Page 8 - exclusion criteria include 'participants missed most of the items in data collection'.

Re: We have checked and modified the inclusion and exclusion criteria on page 7, line 138 to line 146:

The inclusion criteria were: a). All study participants signed the informed consent form before taking part in the survey. b). Participants should complete the physical examination, and should finish at least one of these examination included bone mineral density scan, blood test and questionnaires. And the exclusion criteria were: a). Participants only complete the physical examination, but have no data on bone mineral density scan, blood test or questionnaires. b). Age < 14 year old or > 25 year old. c). Participants have taken drugs which could affect bone metabolism (e.g. glucocorticoids). d). Participants have illness which could cause secondary osteoporosis (e.g. Hyperparathyroidism). In WBBC pilot 3, the urine and faeces of the participants were collected, therefore, participants taking antibiotics should be excluded.

Additionally, the methods are vague in some instances eg. (page 9) - to ensure accurate data participants were asked to have no 'excessive physical activity for at least one night'.

Re: Thanks for your carefully reading, we indeed have wrong statement in this paragraph, we have revised this part accordingly on page 8, line 164 to line 166; page 8, line 175 to line 179. The detail contents are as follows.

To ensure accurate data of resting blood pressure and heart rate, the participants were asked to have rested for at least 5 minutes and have no excessive physical activity.

Participants came to the examination center in each college in the morning with at least 8 h of overnight fasting, about 25 ml of venous blood samples were collected for routine blood measurements, biochemical indexes, DNA extraction and so on. To ensure accurate data, the participants were asked to have tea or alcohol intake or smoking for at least one night before blood sample collection.

There is also no definition for smoking status (just yes or no?) or alcohol consumption (just current drinker or not?).

Re: In WBBC pilot 1, the questionnaire only had “yes” and “no” choices for current smoking and alcohol status, in WBBC pilot 2 and 3, we updated the questionnaire to include the frequency information for smoking and alcohol status. We have added this in the maintext accordingly. Please see page 9, line 199 to line 202.

Figure 1 is particularly difficult to read/interpret and doesn't add much value to the paper. There is no real discussion in the paper at all.

Re: We collected sample from each province from China, and the sample size from each province is different. Figure 1 presents the distribution of samples within China in WBBC pilot cohort. Figure 1A provides a range of sample size in each province with different color, and Figure 1B provides the exact number of sample size in each province. We have made it more clearly in Figure Legend on page 29.

I suggest waiting until longitudinal data provides some scientific novelty, before creating a narrative with clarity in the results, and a strong, separate discussion of the findings.

Re: The WBBC project is an open-ended prospective study, the pilot project of WBBC has focused on the study on the young population (late adolescence). The main purpose of this particular paper is to profile the pilot cohort, and a wide range of phenotype has been collected already. We have set up a website for this pilot cohort (https://wbbc.westlake.edu.cn/index.html), and we have got progress on sequencing data, and the genomic study has been carried out and been described elsewhere (https://www.biorxiv.org/content/10.1101/2021.02.06.430086v1). The future plan of WBBC will collect more samples with old age.

Reviewer: 2
Dr. Janet Olson, Mayo Clinic

Comments to the Author:

1) The WBBC study is a well-designed cohort study with strong likelihood of providing meaningful data in the future. The use of the word ‘pilot’ in the description of this cohort throughout this manuscript is rather confusing. A pilot study implies that you intend to expand this cohort to a larger # of subjects, but this was not discussed anywhere. I recommend dropping the word “pilot” and just refer to the WBBC study. If you are referring to the 3 recruitment phases, perhaps use “phase” rather than “pilot”.

Re: Sorry for the confusion, in fact, Westlake BioBank for Chinese (WBBC) is designed as a large-scale prospective cohort with its aim of recruitinat least 100,000 Chinese samples at different age (https://wbbc.westlake.edu.cn/index.html). The pilot project of WBBC has focused on the study on the

young population (Late adolescence), and has already collected a wide range of information including demographics and anthropometric measures, serological tests, physical activity, sleep quality, age at menarche and bone mineral density etc. We have made it more clearly in the INTRODUCTION. Please see page 4, line 46 to line 59.

2) Providing #s to the 2nd decimal is meaningless. Round to 1 decimal.

Re: Thanks for your suggestion. We have modified the decimal in the maintext and in Tables. Please see ABSTRACT on page 2, Findings to date on page 11 to page 14 and Table 4, Table 5, Table 6 and Table 7. However, the decimal of all data is determined to the accuracy of the equipment, so weight (kg) and grip strength (kg) have the 2nd decimal and bone mineral density (g/cm2) have the 3nd decimal.

3) Use males/females consistently rather than boys/girls, men/women. The latter implies an age stratification that I do not believe the authors intend.

Re: As suggested, we have replaced other words with the words males/females and have revised in the manuscript accordingly (e.g. ABSTRACT on page 2, Findings to date on page 11 to page 14).

4) Minor, but frequent, corrections to English grammar are needed, including referring to adolescence when it should be adolescents, missing plurals, missing articles (a, the). I have included a pdf of my edits/suggestions that are recommendations and not likely exhaustive.

Re: Thanks for your careful reading. We have modified and corrected the minor grammatical errors in the main text. (e.g. page 5, line85, line 90, line 95 and line 196; page 6, line 107 to 108; page 7, line 138, line 149; page 8, line 176; page 9, line 216, line 219 and line 220; page 10, line 226, line 239 and line 243; page 11, line 265, line 271, line278 and line 280; page 12, line 291, line 292, line294 and line 310; page 13, line 332; page 14, line 342.)

5) Inclusion/exclusion criteria should be improved to be more clear. Was it required that all complete all the exams, BMD scans, blood tests and questionnaires? Please clarify. Also, there is no reason to include the inverse of an inclusion criterion as an exclusion criterion.

Re: Thanks for your suggestion, we have updated our inclusion and exclusion criteria, and we avoid to include the inverse of an inclusion criterion as an exclusion criterion. And participants were not required to complete all the exams (page 7, line 138 to line 146).

The inclusion criteria were: a). All study participants signed the informed consent form before taking part in the survey. b). Participants should complete the physical examination, and should finish at least one of other items including bone mineral density scan, blood test and questionnaire. And the exclusion criteria were: a). Age < 14 year old or > 25 year old. b). Participants have taken drugs which could affect bone metabolism (e.g. glucocorticoids). c). Participants have illness which could cause secondary osteoporosis (e.g. Hyperparathyroidism). In WBBC pilot 3, the urine and faeces of the participants were collected, therefore, participants taking antibiotics should be excluded.

6) Line 323 – clarify which phenomenon to which you are referring

Re: The “phenomenon” refers to the high prevalence of underweight in females. To avoid misunderstanding, we now saying “These results might be due to…”, please see page 13, line 322 to line 330. And we have discussed these results.

A silhouette-matching test was administered in mainland China and found that the majority of the female participants indicated a preference to be more slender[1]. Their ideal figure was underweight and was far smaller than the most attractive female figure chosen by male participants[1]. Recently, a study involving 2,023 young female participants (70.5% subjects aged 20-25 years) frm eight Chinese universities[2] showed that 30.55% of the participants were underweight, and 57.39% of them would like to be much thinner, which would lead to more underweight individuals.

7) Lines 360-365 – this is a run-on sentence with multiple ideas. Divide into fewer sentences to clarify and simplify.

Re: Thank you for your advice. We have modified the expression more clearly on page 14, line 363 to line 368:

A questionnaire related to vitamin D and sun exposure was conducted at a university in Nanjing, China and found that 75.0 % of the students lacked sun exposure because they would like to avoid dark skin[3]. In addition, most of the students (82.7 %) used sun protection, and sunscreen use was more popular in females[3]. However, it was reported that using the amount of sun cream recommended by World Health Organization exponentially suppressed vitamin D synthesis in the skin[4].

8) Table 5 – Improve the variable descriptions so it is clear which measures and cutpoints were used to assign subjects to each group.

Re: According to the Working Group on Obesity in China (WGOC)[5], participants were defined as underweight (< 18.5 kg/m2), normal weight (18.5-23.9 kg/m2), overweight (24-27.9 kg/m2) and obese (≥ 28 kg/m2). In addition, central obesity was defined as WC≥85 cm for males and as WC≥80 cm for females based on the recommendations of the WGOC[5]. The detail information is on page 12, line 301 to line 303; line 313 to line 315. We have added this information in the Table Legend.

References

1. Wang K, Liang R, Ma ZL, et al. Body image attitude among Chinese college students. PsyCh journal 2018;7(1):31-40 doi: 10.1002/pchj.200[published Online First: Epub Date]|.

2. Zhang L, Qian H, Fu H. To be thin but not healthy - The body-image dilemma may affect health among female university students in China. PloS one 2018;13(10):e0205282 doi: 10.1371/journal.pone.0205282[published Online First: Epub Date]|.

3. Zhou M, Zhuang W, Yuan Y, et al. Investigation on vitamin D knowledge, attitude and practice of university students in Nanjing, China. Public health nutrition 2016;19(1):78-82 doi: 10.1017/S1368980015000373[published Online First: Epub Date]|.

4. Faurschou A, Beyer DM, Schmedes A, et al. The relation between sunscreen layer thickness and vitamin D production after ultraviolet B exposure: a randomized clinical trial. The British journal of dermatology 2012;167(2):391-5 doi: 10.1111/j.1365-2133.2012.11004.x[published Online First: Epub Date]|.

5. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity in C. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomedical and environmental sciences : BES 2002;15(1):83-96