PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Polymorphisms of genes related to vitamin D metabolism and transportation and its relationship with the risk of osteoporosis: protocol for a multicenter prospective cohort study in China
AUTHORS	wang, jing; Shu, Bing; Li, Chen-guang; Xie, Xing-wen; Liang, De; Chen, Bo-lai; Lin, Xin-chao; Wei, Xu; Wang, Liang; Leng, Xiang-yang; Zhou, Ying-jie; Chen, Peizhan; Tao, Yu-ren; zhou, Yong; Zhang, Yan; Cui, Xue-jun; Lu, Sheng; Wang, Hui; Shi, Qi; Wang, Yong-jun

VERSION 1 – REVIEW

REVIEWER	Alberto Hidalgo-Bravo National Institute of Rehabilitation, Mexico
REVIEW RETURNED	27-Dec-2018

GENERAL COMMENTS	This protocol described the procedures to perform a cohort study to identify association between Bone Mineral Density with genetic and environmental factors. They are proposing to recruit a large population from different regions of China. One concern is the diversity among populations. Are they performing any kind of ancestry test? Are they considering to adjust the analysis by ancestry? They are focusing on the Vitamin D metabolism and transportation and they are planning to measure the Vitamin D metabolites. They should include an item in their questionnaire to investigate sun exposure and month of the year for blood sample collection. Food intake is always difficult to measure are they using standard tables for food equivalents? Vitamin D Is liposoluble, therefore it could be useful to consider body fat content in addition to BMI.

REVIEWER	Neil Gittoes University Hospitals Birmingham UK
REVIEW RETURNED	04-Feb-2019

GENERAL COMMENTS	It is commendable that such a large study has been funded to look at potentially important factors in determining risk of fracture and osteoporosis in a Chinese population. This is a rationale and methods paper and thus by definition, there are no data at this point in time. From the perspective of a clinician with an interest in metabolic bone disease, it is the output of such a study that would be of interest and value but less so from a study design, logistics and methods perspective.

REVIEWER	Kristin Sainani

	Stanford University, United States
REVIEW RETURNED	13-Feb-2019

GENERAL COMMENTS	This is an exciting and important study. But the protocol contains insufficient detail. The stated aim of the study is to determine the relationship between polymorphisms of genes related to Vitamin D and the risk of osteoporosis and osteoporotic fracture. However, though the protocol gives detailed information about the measurement of blood proteins, it gives almost no information about the genetic measurements (e.g., specific SNPs to be measured). Also, no information is given about how osteoporotic fracture will be measured or defined; and more details are needed about the definition of osteoporosis. The statistical plan is also sparse and unclear. Specific comments: 1. The stated main outcomes are osteoporosis and osteoporotic fracture. Is one of these outcomes considered primary? 2. Osteoporosis is defined as T-score <=-2.5. But at what bone site? Is one T-score <=-2.5 (at any bone site, even a single vertebra) sufficient for a diagnosis of osteoporosis? 3. Osteoporotic fracture is listed as one of the main outcomes, but we are given no details about how this will be measured or defined. Is this self-reported fracture? Will these be confirmed by medical records? How will you distinguish between traumatic and osteoporotic fractures? Are you considering all bone sites or only certain sites (e.g., hip and wrist)? 4. The primary exposure of interest is genetic polymorphisms, but there's no information given about the specific SNPs that will be examined. Please provide a Table similar to Table 3 that gives the SNPs (which ones, how they will be measured, how they will be analyzed, etc.). 5. The sample size calculation is unclear. What was the specific statistical test/association that was used in the power calculation--for example, was this for examining the association between 1 SNP or multiple SNPs and osteoporosis or osteoporotic fracture? What was the effect size assumed? Does it consider the fact that many participants may already have osteoporosis at baseline? Was the calculation done for a cross-sectional or longitudinal comparison? 6. The statistical analysis section is too sparse. The authors should divide this section into two: Cross-sectional analyses (for the baseline data) and Longitudinal analyses (for the follow-up data). Many more details are needed about exactly what will be compared and how. For example, currently the cross-sectional comparisons are described with a single sentence: "Multiple logistic regressions will be performed to identify risk factors associated with unfavourable outcomes." This is insufficiently detailed. The goal of the study is to relate polymorphisms to osteoporosis and osteoporotic fracture. Specifically, how will this be done? How will SNPs be handled?Will there will be control for multiple comparisons? For the longitudinal analyses, will those with baseline osteoporosis/osteoporotic fracture be excluded from the analyses? How will missing data be handled? How will model assumptions be tested? What confounders will be considered? Also, the statistical plan appears inappropriate for the stated aims. The aim is not to build a predictive model, but to test the association between genetic polymorphisms and osteoporosis/osteoporitic fracture. The model building process needs to reflect this. The current description "all significant variables in univariate analysis will be used in multivariate analysis" is not an appropriate approach when testing specific hypotheses. 7. Some additional copy editing would be helpful. For example, the

	English in this sentence is unclear: "Secondly, most questions in the questionnaire are retrospective resulting in the deviation, such as medical history and stop or switch medications."

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Alberto Hidalgo-Bravo
Institution and Country: National Institute of Rehabilitation, Mexico
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
This protocol described the procedures to perform a cohort study to identify association between Bone Mineral Density with genetic and environmental factors.
They are proposing to recruit a large population from different regions of China. One concern is the diversity among populations. Are they performing any kind of ancestry test? Are they considering to adjust the analysis by ancestry?

Response: Thank you. We totally agree to concern about the diversity among populations. In order to adjust the difference, we will record the participants' native places, and set it as dummy variable for analysis of covariance (ANCOVA). In addition, we can also observe the distribution of SNPs sites in various regions, and then combine them through meta-analysis.

They are focusing on the Vitamin D metabolism and transportation and they are planning to measure the Vitamin D metabolites. They should include an item in their questionnaire to investigate sun exposure and month of the year for blood sample collection.

Response: Thank you. We totally agree with your instructive suggestions and have revised the Methods section. The date of blood sample collection will be recorded in detail. We also include an item of hours of daily sun exposure in our questionnaires.

Food intake is always difficult to measure are they using standard tables for food equivalents?

Response: Thank you. It is really difficult to evaluate the daily food intakes with standard dietary questionnaires due to the time limit in a large on-site study. So we only focused on the main dietary habits with our own questionnaire, including the frequency and amounts of vegetable oil and animal oil, rice, cooked wheaten food, coarse cereals, salted food, meat, poultry, eggs, seafood, freshwater fish and shrimp, animal viscera, bean products, vegetables, fruits, milk and milk products, tea, coffee, carbonated beverages in a week. It is one of limitations of our study and was discussed in the Strengths and limitations section.

Vitamin D Is liposoluble, therefore it could be useful to consider body fat content in addition to BMI.
Response: Thank you for your suggestions. We have added the measurement of body fat content to the protocol, and the Methods section was revised.

Reviewer: 2
Reviewer Name: Neil Gittoes
Institution and Country: University Hospitals Birmingham - UK
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
It is commendable that such a large study has been funded to look at potentially important factors in determining risk of fracture and osteoporosis in a Chinese population. This is a rationale and methods paper and thus by definition, there are no data at this point in time. From the perspective of a clinician with an interest in metabolic bone disease, it is the output of such a study that would be of interest and value but less so from a study design, logistics and methods perspective.
Response: Thank you for your encouragement and careful reading of our manuscript.

In our previous community based cross-sectional study with 967 postmenopausal women in Shanghai, we found that the variant rs7041 of DBP gene was positively correlated with the total 25(OH)D level but negatively associated with bioavailable 25(OH)D levels. Both total and bioavailable 25(OH)D levels were significantly correlated with the BMD value in postmenopausal women; however, only the bioavailable 25(OH)D level was an independent determinant of the BMD values.

[*Chenguang Li, Peizhan Chen, XiaohuaDuan, et al. Bioavailable 25(OH)D but Not Total 25(OH)D Is an Independent Determinant for Bone Mineral Density in Chinese Postmenopausal Women. EBioMedicine. 2017, 15:184-192.*]

Due to the limits of cross-sectional survey, we could not determine the association of DBP gene polymorphisms, as well as bioavailable 25(OH)D level, with bone fracture risk. To figure it out, we will further perform the prospective cohort study to determine whether DBP gene polymorphisms and bioavailable 25(OH)D level are associated with the fracture risk in both men and women in China. Also, we will observe the distributions of gene polymorphisms of other molecules related to vitamin D metabolism and transportation, and fully characterize the associations of these gene polymorphisms with BMD and fracture risk.

Reviewer: 3
Reviewer Name: Kristin Sainani
Institution and Country: Stanford University, United States
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
This is an exciting and important study. But the protocol contains insufficient detail. The stated aim of the study is to determine the relationship between polymorphisms of genes related to Vitamin D and the risk of osteoporosis and osteoporotic fracture. However, though the protocol gives detailed information about the measurement of blood proteins, it gives almost no information about the genetic measurements (e.g., specific SNPs to be measured). Also, no information is given about how osteoporotic fracture will be measured or defined; and more details are needed about the definition of osteoporosis. The statistical plan is also sparse and unclear. Specific comments:

1. The stated main outcomes are osteoporosis and osteoporotic fracture. Is one of these outcomes considered primary?
Response: Sorry for making confusion. Actually, the major aim of this study is to determine whether the polymorphisms of genes related to Vitamin D have the predictive ability for the BMD loss in the elderly people, and the secondary objective was to determine whether these gene polymorphisms are associated with the risk of osteoporotic fractures.

The description of Outcomes in the Methods section has been revised.

2. Osteoporosis is defined as T-score <=-2.5. But at what bone site? Is one T-score <=-2.5 (at any bone site, even a single vertebra) sufficient for a diagnosis of osteoporosis?

Response: Thank you for your careful reading of our manuscript, and sorry for making confusion. We have revised the description in the Methods section of the manuscript.

According to the suggestion of the Chinese Society of Osteoporosis and Bone Mineral Research states that a clinical diagnosis of osteoporosis can be made in postmenopausal women and men aged 50 and over who sustain a low-trauma fracture (such as hip, vertebra, or radius fracture), or

when the spine and hip BMD are less than or equal to 2.5 standard deviations below the young normal mean (T-score ≤ -2.5) at any bone site, even a single vertebra. For premenopausal women and men under the age of 50, a diagnosis of osteoporosis is established by spine and hip BMD of less than or equal to 2.0 standard deviations below people of the same race, sex and age (Z-score ≤ -2.0) at any bone site, even a single vertebra.

[*Weibo Xia, Zhenlin Zhang, Hua Lin, XiaolanJin, Wei Yu, Qin Fu. Guidelines for the diagnosis and management of primary osteoporosis. Chin J Osteoporo. 2019, 25(3):281-309*]

3. Osteoporotic fracture is listed as one of the main outcomes, but we are given no details about how this will be measured or defined. Is this self-reported fracture? Will these be confirmed by medical records? How will you distinguish between traumatic and osteoporotic fractures? Are you considering all bone sites or only certain sites (e.g., hip and wrist)?

Response: Sorry for making confusion. The information of new fractures are obtained from the participants either in a self-reported manner or by interviews during follow-up visits. Participants with fractures will be asked to describe the causes of fracture (ie, slips, trips or falls; traffic accidents; crushing injury; sharp or blunt trauma and others) and provide their medical records (ie, reports of imaging examinations and treatment). We will defined the fractures (at any bone site) caused by the low-energy injuries (ie, slips, trips and falls) as osteoporotic fractures.

Relative modification was shown in Table 2 (Fracture history) and “Follow-up” section of the manuscript.

4. The primary exposure of interest is genetic polymorphisms, but there's no information given about the specific SNPs that will be examined. Please provide a Table similar to Table 3 that gives the SNPs (which ones, how they will be measured, how they will be analyzed, etc.).

Response: Thank you for your instructive suggestions. The information for about the specific SNPs were provided as Table 4.

Table 4. Single nucleotide polymorphisms of genes

Protein	**Genetic locus of SNPs**
GC	rs4588, rs7041, rs222020, rs2282679
VDR	rs1544410, rs2239181, rs21077301, rs2239179, rs2189480, rs3819545, rs2239186, rs2254210, rs2238136, rs4760648, rs11168287, rs4328262, rs4334089, rs3890733, rs110783219, rs7299460
Vitamin D metabolic enzymes (CYP2R1, CYP27B1)	rs10741657, rs2060793, rs12794714, rs10877012

GC, vitamin D binding protein; ***VDR***, Vitamin D receptor; ***SNPs,*** single nucleotide polymorphisms.

5. The sample size calculation is unclear. What was the specific statistical test/association that was used in the power calculation--for example, was this for examining the association between 1 SNP or multiple SNPs and osteoporosis or osteoporotic fracture? What was the effect size assumed? Does it consider the fact that many participants may already have osteoporosis at baseline? Was the calculation done for a cross-sectional or longitudinal comparison?

Response: Sorry for making confusion.

We estimated the sample size based on the incidence of osteoporosis morbidity in the Japanese population-based osteoporosis study. The prevalence of osteoporosis in women aged 20-79 years old was 36.1% based on total hip BMD [*Ikeda Y, Iki M, Morita A, et al. Age specific values and cutoff levels for the diagnosis of osteoporosis in quantitative ultrasound measurements at the calcaneus with SAHARA in healthy Japanese women: Japanese population-based osteoporosis (JPOS) study. Calcif Tissue Int, 2002, 71:1-9*], whereas in men aged 20-79 years old, the prevalence was 4% [*Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Meta, 2009, 27(5):620-8*].

We used the sample size estimation method of simple random sampling with the formula as follows:

$$n = \frac{u_{\alpha/2}^2 \pi(1-\pi)}{\delta^2}$$

π is the morbidity, and δ is the admissible error. Set δ=0.1, α=0.05, and $u_{\alpha/2}$=1.96 to calculate the sample size. According to the osteoporosis prevalence, for men, π=0.04, n=9218, and for women, π=0.36, n=682.

The minimum requirement is 9218, and we plan to recruit 18000 participants, which is almost 2 times that of the required minimum sample size. On this basis, approximate 3,000 participants will be included in each region.

We have revised the description in the Patient and Public Involvement section of the manuscript.

6. The statistical analysis section is too sparse. The authors should divide this section into two: Cross-sectional analyses (for the baseline data) and Longitudinal analyses (for the follow-up data). Many more details are needed about exactly what will be compared and how. For example, currently the cross-sectional comparisons are described with a single sentence: "Multiple logistic regressions will be performed to identify risk factors associated with unfavourable outcomes." This is insufficiently detailed. The goal of the study is to relate polymorphisms to osteoporosis and osteoporotic fracture. Specifically, how will this be done? How will SNPs be handled?Will there will be control for multiple comparisons? For the longitudinal analyses, will those with baseline osteoporosis/osteoporotic fracture be excluded from the analyses? How will missing data be handled? How will model assumptions be tested? What confounders will be considered? Also, the statistical plan appears inappropriate for the stated aims. The aim is not to build a predictive model, but to test the association between genetic polymorphisms and osteoporosis/osteoporitic fracture. The model building process needs to reflect this. The current description "all significant variables in univariate analysis will be used in multivariate analysis" is not an appropriate approach when testing specific hypotheses.

Response: Thank you. We have revised the statistical analysis section of the manuscript.

"As first step, the characteristics of the study participants (such as age, BMI, et al.) will be shown as the medians and the interquartile range (IQR) or number and proportion. The distributions of the parameters will be shown by mean and standard deviation.

For the baseline data, the normal distribution of parameters (such as biochemical parameters) will be compared between groups using analysis of variance (ANOVA). Comparison of categorical variables (such as genotyping results, smoking, alcohol drinking, comorbidities) will use the χ2 test. Post-hoc testing for the difference between pairs of genotype groups will use Tukey's method to test whether the genotype distribution is consistent with the Hardy–Weinberg equilibrium (HWE). The correlation between the variants will be determined using Pearson's coefficient of correlation. The relationships between BMD and continuous variable (such as biochemical parameters) will be examined using univariate linear models.

In the third step of the analysis, binary regression analyses or multivariate linear regression analyses will be used to identify risk factors associated with unfavorable outcomes.

In the fourth step, a longitudinal assessment of associations and a survival analysis for each incidence of osteoporosis or osteoporotic fracture will used Kaplan-Meier curves in relation to the related risks. Cox multivariate regression models will be used to compare the probability of osteoporosis or osteoporotic fracture in the follow-up cohorts, adjusting for the necessary covariates.

The relative risk (hazards ratios) will be calculated with a 95% confidence interval. The level for statistical significance will be set at α = 0.05 (two-tailed)."

7. Some additional copy editing would be helpful. For example, the English in this sentence is unclear: "Secondly, most questions in the questionnaire are retrospective resulting in the deviation, such as medical history and stop or switch medications."

Response: Thank you for your suggestion. We have revised the expression in my manuscript and ask native English-speaking professionals for help.

VERSION 2 – REVIEW

REVIEWER	Alberto Hidalgo Bravo National Institute of Rehabilitation, Mexico City, Mexico
REVIEW RETURNED	15-May-2019

GENERAL COMMENTS	The authors present a complete study design for the analysis of SNP in Chinese population. My only concern is that they do not mention if the genetic background of their population could be a confusion variable.

REVIEWER	Kristin Sainani Stanford University USA
REVIEW RETURNED	11-May-2019

GENERAL COMMENTS	The authors have done an excellent job on the revision. However, the statistical analysis section remains too vague. "Regression will be performed to identify risk factors associated with unfavorable outcomes" is much too vague. Specific comments: 1. For each analysis proposed (e.g., logistic, Kaplan-Meier, Cox), state specifically which dependent and independent variables will be examined. The primary outcome is listed as the "change in BMD." How will this be calculated? Will BMD be treated as continuous or categorized (e.g., <-2.5/>-2.5)? Also, for the comparison of SNPs, will authors compare homozygotes and heterozygotes to those lacking the SNPs? 2. Please specify if these models will be stratified on any characteristics, in particular gender. 3. Please specify potential confounders that might confound the relationships between SNPs and proteins and BMD/fracture. How will confounding be assessed? 4. Please explain how multiple comparisons will be handled, given that there are a large number of SNPs and proteins being examined. This is a critical point since some SNPs and proteins will almost certainly achieve statistical significance at p<.05 just by chance. 5. Also please organize statistical analysis section into: Evaluation of the primary outcome, Evaluation of the secondary outcome

VERSION 2 – AUTHOR RESPONSE

Reviewer: 3
Reviewer Name: Kristin Sainani
Institution and Country: Stanford University
USA
Please state any competing interests or state ‘None declared’: none declared

Please leave your comments for the authors below
The authors have done an excellent job on the revision. However, the statistical analysis section remains too vague. "Regression will be performed to identify risk factors associated with unfavorable outcomes" is much too vague. Specific comments:
1. For each analysis proposed (e.g., logistic, Kaplan-Meier, Cox), state specifically which dependent and independent variables will be examined. The primary outcome is listed as the "change in BMD." How will this be calculated? Will BMD be treated as continuous or categorized (e.g., <-2.5/>-2.5)? Also, for the comparison of SNPs, will authors compare homozygotes and heterozygotes to those lacking the SNPs?
Response: Thank you for your suggestions. The primary outcome of this study is changes in BMD values for the participants, and “changes in BMD” refers to the differences in BMD values between the yearly follow-up points and the baseline. BMD will be treated as a continuous value. For the comparison of SNPs, all the genotypes will be involved including the homozygotes, heterozygotes and those lacking the SNPs. Pairwise comparison of multiple groups will be performed with least significant difference (LSD) test when homogeneity of variances are satisfied, and Dunnett-t` test for heterogeneity of variance. We have revised the statistical analysis section.

2. Please specify if these models will be stratified on any characteristics, in particular gender.
Response: Thank you for your reminding. We have revised the statistical analysis section. Gender and age will be stratified in these models.

3. Please specify potential confounders that might confound the relationships between SNPs and proteins and BMD/fracture. How will confounding be assessed?
Response: For the multivariate analyses, the best fit of the final model will be selected by the backward step-down method with the Akaike information criterion setting the BMD changes/fracture as the dependent variable and the gender, age, BMI, biochemical parameters, genotype, smoking, alcohol drinking, comorbidities and physical activities as the independent variables. The variation inflation factor (VIF) will be used to determine the multicollinearity problems for the predictor variables. We have revised the statistical analysis section.

4. Please explain how multiple comparisons will be handled, given that there are a large number of SNPs and proteins being examined. This is a critical point since some SNPs and proteins will almost certainly achieve statistical significance at p<.05 just by chance.
Response: Thank you for your suggestions. Pairwise comparison of multiple groups will be performed with least significant difference (LSD) test when homogeneity of variances are satisfied, and Dunnett-t` test for heterogeneity of variance. We have revised the statistical analysis section.

5. Also please organize statistical analysis section into: Evaluation of the primary outcome, Evaluation of the secondary outcome
Response: Thank you for your suggestions. We have revised the statistical analysis section as suggested.

Reviewer: 1
Reviewer Name: Alberto Hidalgo Bravo
Institution and Country: National Institute of Rehabilitation, Mexico City, Mexico
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
The authors present a complete study design for the analysis of SNP in Chinese population. My only concern is that they do not mention if the genetic background of their population could be a confusion variable.
Response: Thank you. We totally agree with you that the genetic background of their population could be a confusion variable. We have modified the strengths and limitations section and the discussion section.

VERSION 3 – REVIEW

REVIEWER	Kristin Sainani Stanford University, USA
REVIEW RETURNED	11-Jul-2019

GENERAL COMMENTS	Thank you for the excellent revisions. A few more changes are needed in the statistical analysis section: 1. The primary aim is to determine the association between gene polymorphisms and BMD changes. But this aim does not match the statistical analysis approach described under "evaluation of the primary outcome." It would not be appropriate to use an automatic selection procedure to build the regression models; rather, model building should center around testing the associations between the

	polymorphisms and BMD. Other variables would be included in the model if they appeared to confound the relationship between the polymorphisms and BMD. Also "BMD changes/fracture" is not the dependent variable for this aim, only BMD changes are. This section should be rewritten; consultation with a statistician may be helpful. 2. What age groups will be used for the stratification on age? This should be pre-specified.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 3
Reviewer Name

Kristin Sainani

Institution and Country

Stanford University, USA

Please state any competing interests or state 'None declared':
None declared
Please leave your comments for the authors below
Thank you for the excellent revisions. A few more changes are needed in the statistical analysis section:
1. The primary aim is to determine the association between gene polymorphisms and BMD changes. But this aim does not match the statistical analysis approach described under "evaluation of the primary outcome." It would not be appropriate to use an automatic selection procedure to build the regression models; rather, model building should center around testing the associations between the polymorphisms and BMD. Other variables would be included in the model if they appeared to confound the relationship between the polymorphisms and BMD. Also "BMD changes/fracture" is not the dependent variable for this aim, only BMD changes are. This section should be rewritten; consultation with a statistician may be helpful.
Response: Thank you for your reminding. We have revised the statistical analysis section. The correlation between the polymorphisms and BMD changes will be determined using Spearman's correlation coefficient.

2. What age groups will be used for the stratification on age? This should be pre-specified.
Response: Thank you for your reminding. We have revised the statistical analysis section. The age will be stratified by every 10 years.