PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Socio-economic status and vitamin D deficiency among women at childbearing age: a population-based case-control study in rural northern China
AUTHORS	Lin, Shiqi; Jiang, Lifang; Zhang, Yuan; Chai, Jian; Li, Jiajia; Song, Xinming; Pei, Lijun

VERSION 1 – REVIEW

REVIEWER	Prof. Dr. Betül Ersoy Celal Bayar University, Faculty of Medicine, Manisa, TURKEY
REVIEW RETURNED	16-Sep-2020

GENERAL COMMENTS	The article was evaluated. In this article, the effect of socioeconomic status on vitamin D levels in women aged 18-40 was investigated. It was concluded that vitamin D deficiency is more common in women with low socioeconomic status and the most affecting factor is low annual income. My comment on the article is below. Introduction -Vitamin D deficiency is said to be common in China. What clinical problems have vitamin D deficiency revealed in previous studies in China? Please present them. Materials and methods -If you randomized the participants, it is interesting that there are equal numbers of participants in three socioeconomic status. -In what season did you get the blood samples? -Which univariate analysis did you use? Please explain - Body mass index is mentioned, how was it calculated? - Some parameters described in the results are not specified in this section. For example, physical exercise, how much has been done, is it outdoors or indoors? Results - There are so many parameters that the reader is confused. Results should be explained more simply Discussion - In the discussion section, it is considered that there is no comment and discussion in accordance with the results. There are more exaggerated interpretations - How could the socioeconomic situation prevent exposure to sunlight? Table 2. - What is BMI grouping made for? - Too many people with chronic illnesses, participants seem to be mostly sick people - This makes the study meaningless. The most important handicap of the study is that most of the selected people have chronic diseases.

REVIEWER	Tom Thacher Mayo Clinic, USA
REVIEW RETURNED	01-Oct-2020

GENERAL COMMENTS	The authors demonstrate an association between measures of SES and vitamin D status in rural women in northern China. They conclude that lower SES was associated with a greater likelihood of vitamin D deficiency. Introduction: Why did the authors choose to study women of childbearing age? The authors should describe the relevance of vitamin D deficiency in this age group. Methods The authors describe the study design as a population-based case-control study. However, they do not describe any population sampling method that would justify the term “population based”. Rather, it appears that the study subjects were volunteers, but no description of subject recruitment was provided. There is no description of a case-control design for the study. They refer to the Study on Population-based Birth Defects Monitoring. A reference should be provided for this study. What are the latitudes of the 4 counties in China? Data collection: What chronic diseases were included? How were dietary intakes measured? What behavioral exposure factors were assessed? How was utilization of public health services defined? Was data collected on vitamin D supplement use? Are any foods or other vehicles routinely fortified with vitamin D in China? Was there any standardization of 25(OH)D measurements? Did the laboratory participate in a vitamin D standardization program? Socioeconomic status: How was “annual income was enough for expenditure” assessed? Results Table 1: I would be interested in also including the category of 25(OH)D values <12 ng/mL as this is a level clearly associated with risk of adverse outcomes. Table 2: How was “picky eating habits” defined? Why only passive smoking but no personal smoking? How were physical exercise, accept eugenic publicity, and accept physical examination during the past year defined? These operational definitions should be included in the Methods. I believe Table 2 would be more useful if mean±SD 25(OH)D values were given for each of the exposure variable categories, instead of the columns with proportions of “sufficient,” “insufficient,” and “deficient.” Categories for BMI should be the standard ones of <25, 25 to <30, 30 or greater. Logistic regression analysis (Table 3): Why were only statistically significant variables included in the logistic regression models? Was there any assessment for collinearity of variables in the models? For example, I would think that household annual income would be highly correlated with household income for expenditure and other SES variables. I am surprised that Table 4 indicates that 89% of women had chronic diseases. What kinds of chronic diseases were prevalent in women of childbearing age? Discussion The authors mention that the prevalence of severe vitamin D deficiency (<10 ng/mL) was 15.99%. It would be valuable to include this category in the Results section. The authors should compare their proposed SES measure with other SES measures used in studies of vitamin D status.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author

The article was evaluated. In this article, the effect of socioeconomic status on vitamin D levels in women aged 18-40 was investigated. It was concluded that vitamin D deficiency is more common in women with low socioeconomic status and the most affecting factor is low annual income. My comment on the article is below.

Introduction

-Vitamin D deficiency is said to be common in China. What clinical problems have vitamin D deficiency revealed in previous studies in China? Please present them.

Answer: In China, it's reported that vitamin D deficiency/insufficiency is pertinent to clinical issues besides skeletal problems, such as metabolic syndrome[1] and its complications[2], dyslipidemia[3] and cardiovascular diseases[4] and even emotional, behavioral and attentional problems [5], depression[6], as well as reduced sperm quality [7] and lower total testosterone[8].

This part has been added in introduction on Page 3 of the main document.

Materials and methods

-If you randomized the participants, it is interesting that there are equal numbers of participants in three socioeconomic status.

Answer: In our study, we constructed SES score through principal component analysis and then divided the score into 3 categories with the 33rd and 66th percentiles as the cut-off points. So in principle, there should be equal numbers of participants, that is, 383 in each group. Yet since some individuals have equal scores and there were some missing socio-economic status data of individuals, we could only made it 372 in the high SES group, 383 in the middle and 387 in the lowest group.

-In what season did you get the blood samples?

Answer: The blood samples were collected in winter, namely, December 2009, January and February 2010. It was made clear in the Method part on Page 5 of the main document. Therefore, as discussed in limitation part, the level of serum 25(OH)D might be underestimated because the blood sample of this study was collected when daylight and temperature was lower than that in summer.

-Which univariate analysis did you use? Please explain

Answer: We took χ^2 test in comparing vitamin D status across different groups of characteristics of study participants. And it has been added in the method part on Page 8 of the main document.

- Body mass index is mentioned, how was it calculated?

Answer: The body mass index (BMI) was calculated by weight/height^2. The unit is kilograms for weight and meters for height. The height and weight of participants were got through physical examination by trained health workers during investigation. It's made clear on Page 7 of the main document.

- Some parameters described in the results are not specified in this section. For example, physical exercise, how much has been done, is it outdoors or indoors?

Answer: All parameter descriptions were added to the Method part on Page 7-8 of the main document under the subtitle “Covariates”. Taking physical exercises meant taking any one of the following indoor or outdoor exercises at least once a week for more than 30 minutes per time: walking, running, ball games, Tai Chi or other health-promotion physical exercises, swimming and other sports. The question did not distinguish indoor or outdoor activities and thus couldn’t reflect sunlight exposure time, which was also one limitations of our study and was discussed in limitation part.

Results

- There are so many parameters that the reader is confused. Results should be explained more simply

Answer: The result part has been made simpler by leaving out unimportant and insignificant parameters on Page 9-10 of the main document.

Discussion

- In the discussion section, it is considered that there is no comment and discussion in accordance with the results. There are more exaggerated interpretations

Answer: We reexamined the discussion part and tried to made it more confined to results on Page 10-14 of the main document.

- How could the socioeconomic situation prevent exposure to sunlight?

Answer: People from different socioeconomic situation might have different tendencies to sunlight exposure for social or cultural reasons, which has also been found in some previous studies[9-11]. On one hand, people in higher socioeconomic status might be less exposed to sunlight. One reason might be need of work. Farmers or construction workers, for examples, almost can’t avoid being exposed to long-term sunlight while white-collars perform their work mostly indoors. Another reason could be women of higher socioeconomic status, under the current popular culture that adores white skin as beauty, tend to avoid sunlight exposure. Studies did find that females and individuals of a younger age and higher education level were more likely to perform sun-protective behaviors[12]. In fact, in our data, using cosmetics is also associated with vitamin D deficiency. On the other hand, people with health literacy might understand the importance of getting sunlight exposure in gaining vitamin D, which is also reported by some studies[13]. In a word, there might exist complex association between socioeconomic status and sunlight exposure. For lack of data, we can’t explore in detail this relationship but could only raise a possibility in our paper.

Table 2.

- What is BMI grouping made for?

Answer: Studies have shown that obesity is corelated with vitamin D insufficiency or deficiency[14] so here we used body mass index (BMI) to measure obesity. Though BMI group <25, 25 to <30, 30 or greater is most often used in literature, the “<24 as normal weight or underweight, 24 to <28 as overweight and ≥28 as obesity” standard, is the most widely used Chinese standard[15] and might also be the most suitable one for Chinese population to identify obesity-related diseases[16].

- Too many people with chronic illnesses, participants seem to be mostly sick people

Answer: I should apologize that this is the clerical error that I mistakenly inverted the “yes” and “no” in table 2. The “history of chronic diseases” meant to have been diagnosed of any one of the following diseases: anemia, hypertension, hyperlipemia, heart disease, diabetes, hyperglycemia, thyroid diseases, phenylketonuria, epilepsy, asthma, chronic renal diseases, systematic lupus erythematosus, rheumatic arthritis, deep vein thrombosis, cancer, depression or anxiety and schizophrenia. There were 106 (9.2%) participants reported to have had anemia, 12 (1%) for hypertension, 4 (0.3%) for hyperlipemia, 3 (0.3%) for heart diseases, 2 (0.2%) for diabetes, 1 (0.1%) for hyperglycemia, 5 (0.4%) for thyroid diseases, 0 (0%) for phenylketonuria, 2 (0.2%) for epilepsy, 0 (0%) for asthma, 0 (0%) for chronic renal diseases, 1 (1%) for systematic lupus erythematosus, 2 (0.2%) for rheumatic arthritis, 0 (0%) for deep vein thrombosis, 1 (0.1%) for cancer, 4 (0.3%) for depression or anxiety, 3 (0.3%) for schizophrenia. Altogether, 130 (11.29%) participants once had at least one of the above diseases. I’ve already revised the table on Page 18 of the main document.

- This makes the study meaningless.

The most important handicap of the study is that most of the selected people have chronic diseases.

Reviewer: 2

Comments to the Author

The authors demonstrate an association between measures of SES and vitamin D status in rural women in northern China. They conclude that lower SES was associated with a greater likelihood of vitamin D deficiency.

Introduction: Why did the authors choose to study women of childbearing age? The authors should describe the relevance of vitamin D deficiency in this age group.

Answer: Maternal vitamin D deficiency/insufficiency has been found to be associated not only with adverse gestational and neonatal outcomes like low birth weight, prematurity[17], and gestational diabetes mellitus[18], but also with offspring vitamin D deficiency[17], impaired intra-uterine growth[29-30], type I diabetes, nutritional rickets and pneumonia in adulthood[19]. Albeit still inconclusive, maternal vitamin D might begin its vital role in fetal development in early pregnancy[29], thus suggesting adequate vitamin D concentration when preparing for pregnancy. In a word, given that sufficient vitamin status is so vital for women themselves and their offspring and that serum 25OHD concentrations were low among them[3] and their newborn babies in Chinese populations[20-22], it’s of significance to target the vitamin D status and its influencing factors of women at childbearing age. This part has been added in Introduction part on Page 3-4 of the main document.

Methods

The authors describe the study design as a population-based case-control study. However, they do not describe any population sampling method that would justify the term “population based”. Rather, it appears that the study subjects were volunteers, but no description of subject recruitment was provided. There is no description of a case-control design for the study. They refer to the Study on Population-based Birth Defects Monitoring. A reference should be provided for this study.

Answer: The data of the present research was based on the Study on Population-based Birth Defects Monitoring and Comprehensive Intervention Project which aimed to establish a prospective cohort of married but unpregnant women at child-bearing age in 2008-2009 in Henan Province, collect their

baseline characteristics including basic demographic and socio-economic characteristics, dietary intake and behavioral factors as well as their blood samples, follow them until their pregnancy results were observed so as to explore the association between pre-pregnancy risk factors and pregnancy results. A multi-stage cluster sampling method was used to obtain a representative sample of targeted population. 1) In the first stage, 4 counties (Hui County, Mengzhou County, Xinmi County and Luanchuan County) were randomly selected from 158 counties in Henan Province; 2) In the second stage, 40 towns (the next administrative unit below County) were randomly selected from the 4 counties. 3）In the third stage, 5 villages were randomly selected from each town. 4) In the fourth stage, 10 women at child-bearing age and their husbands were randomly selected in each village. The selection criteria was 1) married women with local permanent residency,2) between 18-40 years old, 3) not pregnancy at present, 4) living in the research counties with local registered permanent residency and 5) without any severe heart, liver, kidney, metabolic diseases, blood or other system diseases or cancers. Finally, 1151 of 2000 women had pregnancy results and were thus included in the project. In our study, cases were defined as women with serum 25(OH)D <30 ng/mL and were further subdivided into three groups: 1) vitamin D insufficiency: serum 25(OH)D ≥20–and <30 ng/ml; 2) vitamin D deficiency: serum 25(OH)D ≥10 and <20 ng/ml; 3) Severely deficiency: serum 25(OH)D<10ng/ml. Controls were those with serum 25(OH)D ≥30 ng/mL. Altogether, there were 369 controls and 782 cases, among whom 232 were of vitamin D insufficiency, 366 of deficiency and 184 of severely deficiency.

Since no English-language papers were published based on this project before, a reference was not available. All of the above was added in the Method part under the subtitle “Study design and population” on Page 4 of the main document.

What are the latitudes of the 4 counties in China?

Answer: The latitudes of the 4 counties were 35°17′N (Hui County), 34°50’ N (Mengzhou County), 34°32' N (Xinmi County) and 35°51’ N (Luanchuan County). It’s made clear in the Method Part on Line 5-6, Page 5.

Data collection: What chronic diseases were included? How were dietary intakes measured? What behavioral exposure factors were assessed? How was utilization of public health services defined?

Was data collected on vitamin D supplement use? Are any foods or other vehicles routinely fortified with vitamin D in China?

Answer: Chronic diseases included: anemia, hypertension, hyperlipemia, heart disease, diabetes, hyperglycemia, thyroid diseases, phenylketonuria, epilepsy, asthma, chronic renal diseases, systematic lupus erythematosus, rheumatic arthritis, deep vein thrombosis, cancer, depression or anxiety, and schizophrenia. Nutritional supplement was evaluated by having taken any of the following during the past month: vitamin A, multivitamin B, vitamin B1, vitamin B2, vitamin B6, vitamin B12, vitamin C, vitamin E, cod-liver oil or vitamin D, iron preparations, calcium tablets and zinc supplements. Food intake was measured by average frequencies of the food intake during the past year. The frequencies included “everyday”, “4-6 times per week”, 1-3 times per week”, “1-3 times per month” and “hardly ever” and were divided into 2-3 groups according to distribution of the answered frequencies of different food.

Behavioral factors included picky eating habits, passive smoking and physical exercises. Picky eating habits were measured by the question “do you have picking eating habits, i.e., having preferences to some special food such as fruit and vegetables and keeping eating them every day while rejecting other kinds of food such as meat?”; having passive smoking was defined as “being passively inhaled smoke by smokers around you for more than 15 minutes every day”; taking physical exercises meant taking any one of the following at least once a week for more than 30 minutes per time: walking,

running, ball games, Tai Chi or other health-promotion physical exercises, swimming and other sports.

Utilization of health services included accepting eugenic publicity and accepting physical examination during the past year. Having accepted eugenic publicity was defined as having received materials about knowledge of eugenics such as how to prepare for pregnancy from health service institutions during the past year; having accepted physical examination during the past year was defined as having received systematic inspections of the body for signs and symptoms of disease or abnormality during the past year.

All of the above was added in the Method part on Page 6-8 of the main document under the subtitle "Definitions of socio-economic status" and "Covariates".

There is no foods or other vehicles routinely fortified with vitamin D in China.

Was there any standardization of 25(OH)D measurements? Did the laboratory participate in a vitamin D standardization program?

Answer: The measurement of serum 25 (OH)D was performed by Beijing Mass Spectrometry Medical Research Ltd. (Beijing, China), a well-recognized independent institution for clinical diagnostics. Since there were no general standard of 25(OH)D measurements currently in China, the laboratory didn't participate in any vitamin D standardization program or use any standardization of 25(OH)D measurements. Admittedly, this might affect comparison between the present study and other studies and this limitation was also added in Discussion part on Lin 6-10, Page 14 and Strengths and Limitation Highlights.

Socioeconomic status: How was "annual income was enough for expenditure" assessed?

Answer: Household income for expenditure was measured by the question "whether your family have enough income for expenditure in your daily life?" and we grouped the answers of "a lot more income than expenditure", "a little more income than expenditure" and "balanced" into "surplus" and "income is not enough for expenditure" and "a lot more expenditure than income" into "inadequate or deficit". It was made clear in the Method part on Line 7-13 from the bottom of Page 6.

Results

Table 1: I would be interested in also including the category of 25(OH)D values <12 ng/mL as this is a level clearly associated with risk of adverse outcomes.

Answer: The category of 25(OH)D<10ng/ml has already been added in analysis.

Table 2: How was "picky eating habits" defined? Why only passive smoking but no personal smoking? How were physical exercise, accept eugenic publicity, and accept physical examination during the past year defined? These operational definitions should be included in the Methods.

Answer: Picky eating habits were measured by the question "do you have picking eating habits, i.e., having preferences to some special food such as fruit and vegetables and keeping eating them every day while rejecting other kinds of food such as meat?"; having passive smoking was defined as "being passively inhaled cigarette smoke by smokers around you for more than 15 minutes every day". Having accepted eugenic publicity was defined as having received materials about knowledge of eugenics such as how to prepare for pregnancy from health service institutions during the past year; having accepted physical examination during the past year was defined as having received systematic inspections of the body for signs and symptoms of disease or abnormality during the past year. All of the above was added in the Method part on Page 7-8 of the main document under the subtitle "Covariates".

We did have questions about personal smoking but only 4(0.3%) answered “yes”. Considering common situations in real lives that women seldom smoke in China and reports of other studies that the prevalence of tobacco smoking is traditionally low among Chinese women, about 6.4% in 2003[23], the figure in our study is reasonable. Therefore, we didn’t include personal smoking here. Yet since non-smokers might also be exposed to tobacco smokes and passive smoking has also been found to be associated with vitamin D deficiency [24], we included passive smoking here.

I believe Table 2 would be more useful if mean±SD 25(OH)D values were given for each of the exposure variable categories, instead of the columns with proportions of “sufficient,” “insufficient,” and “deficient.” Categories for BMI should be the standard ones of <25, 25 to <30, 30 or greater.

Answer: The table 2 has been revised by giving counts and percentages and medium (25^{th} ~75^{th} percentiles) for each of the exposure variable categories given that vitamin D level in our study was not normally distributed. The proportions of “sufficient,” “insufficient”, “deficient” and “severely deficient” were still reserved because if absolute numbers between cases and controls across groups had not been shown, we are afraid that readers couldn’t see whether there were enough test power in multivariate analysis. For the conciseness of the table, the original P value column were dropped and a subscript “*” were used to denote significant P values of χ^2 test.

Though BMI group <25, 25 to <30, 30 or greater is most often used in literature, the “<24 as normal weight or underweight, 24 to <28 as overweight and ≥28 as obesity” standard is the most widely adopted Chinese standard[15] and might also be the most suitable one for Chinese population to identify obesity-related diseases[16]. Besides, in our study, 879 (76.4%) participants’ BMI were <25, 215 (18.7%) were 25 to <30, and only 54 (4.7%) were 30 or greater. The number in “30 or greater” group was so small that when was further divided into three groups of vitamin D status, the number might be too small to make comparison meaningful. Therefore, we still keep 24 and 28 as cut-off points here.

Logistic regression analysis (Table 3): Why were only statistically significant variables included in the logistic regression models? Was there any assessment for collinearity of variables in the models? For example, I would think that household annual income would be highly correlated with household income for expenditure and other SES variables.

Answer: We conducted univariate analysis and put significant variables into multivariate models in order to find out confounding factors that should be controlled when exploring the association between SES and vitamin D status. Though in observational studies, confounders are untestable[25], factors that are not confounders could possibly be detected considering that, if confounders exist differences between outcome groups, then factors that don’t have differences between outcome groups can be seen as non-confounders. This method could help make the final multivariate model more succinct, reduce collinearity that too many parameters might bring about and increase the test power. Yet this method was not without its limitations[26] and admittedly, the significance level of 0.05 was made kind of by arbitrary, and possibilities were that insignificant variables in univariate analysis could become significant in multivariate variables. Therefore, we revised our results by including all variables into multivariate models. Please kindly see table 3 and 4 in detail. As can be seen, the results remained similar, reflecting our results relatively robust to some extent.

We admitted that there might exist problems of collinearity considering so many parameters in the model. Especially different dimensions of SES would correlate with each other and that’s one of the reasons we further construct a composite SES index. To test the severity of collinearity, we conducted multicollinearity analysis through calculating VIF of different models respectively and the results are listed in the following table:

Supplementary table 3. The Variance Inflation Factor for the Multivariate Models

Variables	Insufficiency model	Deficiency model	Severe deficiency model
Age	1.29	1.33	1.33
BMI	1.06	1.05	1.06
Gravidity	1.33	1.05	1.34
History of chronic diseases	1.06	1.03	1.04
Women's education	1.33	1.28	1.30
Husband's education	1.27	1.22	1.24
Women's occupation	1.24	1.21	1.21
Husband's occupation	1.09	1.10	1.11
Household annual income	1.20	1.13	1.19
Household income for expenditure	1.38	1.25	1.27
Nutritional supplement	1.20	1.14	1.17
Eggs intake	1.24	1.25	1.22
Meat intake	1.16	1.16	1.14
Fish intake	1.23	1.25	1.26
Milk intake	1.36	1.38	1.42
Bean intake	1.30	1.35	1.31
Vegetables intake	1.17	1.20	1.12
Picky eating habits	1.09	1.04	1.09
Passive smoking	1.05	1.04	1.04
Physical exercise	1.12	1.12	1.16
Accept eugenic publicity	1.28	1.29	1.27
Physical examination	1.18	1.20	1.27

According to the table above, albeit multicollinearity did exist, all VIF remained between 1.01 ~ 1.38, far below the commonly recognized multicollinearity threshold 10. Therefore, we consider the collinearity might not be a big issue in our models.

I am surprised that Table 4 indicates that 89% of women had chronic diseases. What kinds of chronic diseases were prevalent in women of childbearing age?

Answer: I should apologize that this is the clerical error that I mistakenly inverted the "yes" and "no" in table 2. There were 106 (9.2%) participants reported to have had anemia, 12 (1%) for hypertension, 4 (0.3%) for hyperlipemia, 3 (0.3%) for heart diseases, 2 (0.2%) for diabetes, 1 (0.1%) for hyperglycemia, 5 (0.4%) for thyroid diseases, 0 (0%) for phenylketonuria, 2 (0.2%) for epilepsy, 0

(0%) for asthma, 0 (0%) for chronic renal diseases, 1 (1%) for systematic lupus erythematosus, 2 (0.2%) for rheumatic arthritis, 0 (0%) for deep vein thrombosis, 1 (0.1%) for cancer, 4 (0.3%) for depression or anxiety, and 3 (0.3%) for schizophrenia. Altogether, 130 (11.29%) participants once had at least one of the above diseases. I've already revised the table 1.

Discussion

The authors mention that the prevalence of severe vitamin D deficiency (<10 ng/mL) was 15.99%. It would be valuable to include this category in the Results section.

Answer: The severe deficiency group has already been included in the Results section.

The authors should compare their proposed SES measure with other SES measures used in studies of vitamin D status.

Answer: There are different ways of measuring SES in studies of vitamin D status. The most commonly seen is focusing on one dimension of SES by using individual SES indicators such as education attainment, income, expense management, and occupation[27, 28]. Indicators such as 'poverty-income ratio', a ratio of family income to poverty threshold used in National Center for Health Statistics conducts the National Health and Nutrition Examination Survey (NHANES) in US[29], also focuses on one dimension of SES, i.e., family income, but takes into consideration of local development to create an index[30]. In developing composite SES indexes, scales were given to different indicators or indicator combinations. For example, Kuppuswamy's socioeconomic status, a relatively well-established tool in India, assigns 7 scores to education, 10 for occupation, and 12 for family income and makes 5 groups of SES [31] and SES index for German Health Interview and Examination Survey for Adults (DEGS1), assigns different score to different combination of specific kinds of occupation, income and education[32]. While some others develop a specific questionnaire based on local conditions to get an SES score. For instance, EPICES, a French evaluation of low socio-economic status and inequalities in Health Examination Centers, aggregates a lot more social dimensions such as leisure activities[33]. In our study, we not only evaluated several individual dimensions of SES to examine their separate associations to vitamin D insufficiency/deficiency, but also constructed an SES index. Since there is no well-recognized way of measuring SES in local area and it seems lack of credibility to assign scores to different SES dimensions by arbitrary, we used a 'data-driven method' through the principal component analysis to developed the index. Though hardly seen in analyzing vitamin D status, principal component analysis has proven to be quite validated and robust in constructing SES in other epidemiological studies[1]. By summing effects of individual SES indicators, we could get a better view of association between SES and vitamin D insufficiency/deficiency and increase the test power.

The above has been added to the Discussion part on Page 11-12.

References

1. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning* 2006;21(6):459-68.
2. Xiao Y, Wei L, Xiong XF, et al. Association Between Vitamin D Status and Diabetic Complications in Patients With Type 2 Diabetes Mellitus: A Cross-Sectional Study in Hunan China. *Front Endocrinol* 2020;11:11. doi: 10.3389/fendo.2020.564738
3. Jiang XJ, Peng M, Chen SH, et al. Vitamin D deficiency is associated with dyslipidemia: a cross-sectional study in 3788 subjects. *Curr Med Res Opin* 2019;35(6):1059-63. doi: 10.1080/03007995.2018.1552849
4. Wang T, Sun HL, Ge HN, et al. Association between vitamin D and risk of cardiovascular disease in Chinese rural population. *Plos One* 2019;14(5):12. doi: 10.1371/journal.pone.0217311
5. Wang XR, Jiao XT, Xu MQ, et al. Effects of circulating vitamin D concentrations on emotion, behavior

and attention: A cross-sectional study in preschool children with follow-up behavior experiments in juvenile mice. *J Affect Disord* 2020;275:290-98. doi: 10.1016/j.jad.2020.06.043
6. Yao Y, Fu SH, Zhang H, et al. The prevalence of depressive symptoms in Chinese longevous persons and its correlation with vitamin D status. *BMC Geriatr* 2018;18:7. doi: 10.1186/s12877-018-0886-0
7. Chen YL, Liu DF, Zeng L, et al. Effect of serum 25-hydroxyvitamin D levels on sperm quality and assisted reproductive technology outcomes for men of infertile Chinese couples. *Andrology* 2020;8(5):1277-86. doi: 10.1111/andr.12818
8. Chen C, Zhai HL, Cheng J, et al. Causal Link Between Vitamin D and Total Testosterone in Men: A Mendelian Randomization Analysis. *J Clin Endocrinol Metab* 2019;104(8):3148-56. doi: 10.1210/jc.2018-01874
9. Hunt-Watts HJ, Cade JE, Hadley DM. Food and nutrient intake in low-income families: the archaeology of nutrition. *Proceedings of the Nutrition Society* 2015;74(OCE5):E358-E58. doi: 10.1017/s002966511500405x
10. Narchi H. Case-control study of diet and sun exposure in adolescents with symptomatic rickets. *Ann Trop Paediatr* 2000;20(3):217-21. doi: 10.1080/02724936.2000.11748137
11. Hansen MR, Bentzen J. High-risk sun-tanning behaviour: a quantitative study in Denmark, 2008-2011. *Public Health* 2014;128(9):777-83. doi: 10.1016/j.puhe.2014.07.002
12. Yan S, Xu F, Yang C, et al. Demographic Differences in Sun Protection Beliefs and Behavior: A Community-Based Study in Shanghai, China. *International Journal of Environmental Research and Public Health* 2015;12(3):3232-45. doi: 10.3390/ijerph120303232
13. Leung AYM, Cheung MKT, Chi I. Supplementing vitamin D through sunlight: Associating health literacy with sunlight exposure behavior. *Arch Gerontol Geriatr* 2015;60(1):134-41. doi: 10.1016/j.archger.2014.10.005
14. Jang H, Lee Y, Park K. Obesity and Vitamin D Insufficiency among Adolescent Girls and Young Adult Women from Korea. *Nutrients* 2019;11(12):8. doi: 10.3390/nu11123049
15. Qin Y, Melse-Boonstra A, Pan X, et al. Anemia in relation to body mass index and waist circumference among chinese women. *Nutrition Journal* 2013;12 doi: 10.1186/1475-2891-12-10
16. Zhou F. Cooperative Meta-Analysis Group of the Working Group on Obesity in China 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;15:83-96.
17. Wang Y, Li H, Zheng M, et al. Maternal vitamin D deficiency increases the risk of adverse neonatal outcomes in the Chinese population: A prospective cohort study. *PLoS One* 2018;13(4):e0195700. doi: 10.1371/journal.pone.0195700 [published Online First: 2018/04/25]
18. Al-Shafei AI, Rayis DA, Mohieldein AH, et al. Maternal early pregnancy serum level of 25-Hydroxyvitamin D and risk of gestational diabetes mellitus. *Int J Gynecol Obstet*:4. doi: 10.1002/ijgo.13389
19. Wang S, Shen G, Jiang S, et al. Nutrient Status of Vitamin D among Chinese Children. *Nutrients* 2017;9(4):319. doi: 10.3390/nu9040319
20. Wang J, Zhang JY, Wei R, et al. Serum 25 hydroxyvitamin D status in 6-month-old infants in Guangzhou, China: A paired longitudinal follow up study. *Matern Child Nutr* 2020;16(2):9. doi: 10.1111/mcn.12924
21. Wei F, Wang Z, Wang JJ, et al. Serum vitamin D levels among children aged 0-12 years in the First Affiliated Hospital of Harbin Medical University, China. *J Public Health* 2018;40(4):721-26. doi: 10.1093/pubmed/fdy055
22. Zhang W, Stoecklin E, Eggersdorfer M. A glimpse of vitamin D status in Mainland China. *Nutrition* 2013;29(7):953-57. doi: https://doi.org/10.1016/j.nut.2013.01.010
23. ZHANG J, OU J-X, BAI C-X. Tobacco smoking in China: Prevalence, disease burden, challenges and future strategies. *Respirology* 2011;16(8):1165-72. doi: https://doi.org/10.1111/j.1440-1843.2011.02062.x
24. Banihosseini SZ, Baheiraei A, Shirzad N, et al. The effect of cigarette smoke exposure on vitamin D level and biochemical parameters of mothers and neonates. *Journal of diabetes and metabolic disorders* 2013;12(1):19-19. doi: 10.1186/2251-6581-12-19
25. Geng Z, Guo JH, Fung WK. Criteria for confounders in epidemiological studies. Journal of the Royal Statistical Society Series B-Statistical Methodology. 2002;64:3-15.
26. Howards PP. An overview of confounding. Part 2: how to identify it and special situations. Acta Obstetricia Et Gynecologica Scandinavica. 2018;97(4):400-406.
27. Kim JS. Factors Associated with Vitamin D Status Among Korean Female Adolescents. *Journal of*

Pediatric Nursing-Nursing Care of Children & Families 2019;44:E79-E83. doi: 10.1016/j.pedn.2018.11.005
28. Wyskida M, Owczarek A, Szybalska A, et al. Socio-economic determinants of vitamin D deficiency in the older Polish population: results from the PolSenior study. *Public Health Nutrition* 2018;21(11):1995-2003. doi: 10.1017/s1368980017003901
29. Centers for Disease Control and Prevention (CDC) NCfHSN. National health and nutrition examination survey data. In: Hyattsville MUDoHaHS, ed., 2009.
30. Ginde AA, Sullivan AF, Mansbach JM, et al. Vitamin D insufficiency in pregnant and nonpregnant women of childbearing age in the United States. *American Journal of Obstetrics and Gynecology* 2010;202(5):436.e1-36.e8. doi: 10.1016/j.ajog.2009.11.036
31. Sathish P, Sajeethakumari R, Ramasamy P, et al. Correlation between maternal and neonatal blood vitamin D levels and its effect on the newborn anthropometry. *International Journal of Reproduction, Contraception, Obstetrics and Gynecology* 2016 doi: 10.18203/2320-1770.ijrcog20162821
32. Rabenberg M, Scheidt-Nave C, Busch MA, et al. Vitamin D status among adults in Germany - results from the German Health Interview and Examination Survey for Adults (DEGS1). *Bmc Public Health* 2015;15:15. doi: 10.1186/s12889-015-2016-7
33. Leger-Guist'hau J, Domingues-Faria C, Miolanne M, et al. Low socio-economic status is a newly identified independent risk factor for poor vitamin D status in severely obese adults. *Journal of Human Nutrition and Dietetics* 2017;30(2):203-15. doi: 10.1111/jhn.12405

VERSION 2 – REVIEW

REVIEWER	Tom Thacher Mayo Clinic, Rochester, MN, USA
REVIEW RETURNED	11-Jan-2021

GENERAL COMMENTS	The authors have adequately addressed my comments. The reporting of methodology is more complete and robust.