PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The concordance of chronic conditions among the household members in Shanghai: a cross-sectional study
AUTHORS	Qin, Yingyi; Guo, Yibin; Tang, Yuanjun; Wu, cheng; Zhang, Xinji; He, Qian; He, Jia

VERSION 1 – REVIEW

REVIEWER	Ian Fyffe Simon Fraser University, Canada
REVIEW RETURNED	23-Jul-2019

GENERAL COMMENTS	There is very little that I would suggest changing to this manuscript. One is changing the word "scene" to "sub-sample". Another is expanding the limitations section to include how household cohabitation in this instance includes health behaviours such as diet, exercise and second hand smoke as well as biological factors such as genetics. Although this was addressed earlier in the manuscript, I think it would be important to reiterate since the study is ultimately unable to pinpoint the underlying causes.

REVIEWER	Danilo Silva Federal University of Sergipe – Brazil
REVIEW RETURNED	23-Sep-2019

GENERAL COMMENTS	Major points Point 1: The study investigates an interesting issue regarding the co-occurrence of similar heath outcome in people living in the same household. Besides large sample size, my concern is whether sampling process was in accordance with the aim of the current study (more than one person per household). If not, I think some bias is possible. For instance, what differentiate houses that have more than one respondent? Point 2: The self-reported measures are understandable given the number of participants. However, mainly for chronic diseases, some people can have the disease but do not have the diagnostic. For the current study, this could bring a perspective of the diagnosis and health care. When a member of the household receive the diagnostic of some chronic condition, are the other more prone to know and to perform diagnostic tests? Point 3: Are covariates actually moderators/mediators? I think the discussion should consider this. If people in the same household cluster diseases, this could be explained by some of the covariates or other not controlled such as diet and physical activity. Thus, I suggest study these exposures factors instead control them. This

	could also help the understanding about the discussion "environment vs genetic". Minor point Point 1: Information on China's National Health Service Survey in the last paragraph of the introduction (page 5) should be referenced.

REVIEWER	Grant Ritter Brandeis. University Massachusetts, USA
REVIEW RETURNED	30-Sep-2019

GENERAL COMMENTS	General Comments: The research question is interesting and important, and the methods used in the analyses appear sound. The study is worth publishing. However, the current writing is not up to the usual level for a journal article and would benefit from further work. While many readers may be able to figure out much of the paper's content (particularly, if they study the accompanied tables and read the STROBE in the appendices), it needs another round of editing. Hopefully, the authors can be helped by someone with more experience in this area. Grammatical problems (e.g., verb tenses, subject-verb disagreement, dangling or misplaced modifiers, and sentence fragments) should be corrected, and the editor should also revise the content and sequence of written passages to provide better flow and improve readability. In particular, a full rewrite of the analysis plan - what the authors wish to test and how they plan to test it - would be beneficial. The paper uses a number of phrases to describe the analyses ('explored the relationships', 'estimate the risk of chronic conditions when family members has (sic) one or more chronic conditions', 'estimation of the association between one's chronic condition status and their household members' etc.), which are generally too vague. In the Methods section when the GEE logistical model is finally introduced, details are lacking regarding model specification and the composition of the sample and subsamples used to test various hypotheses. With revision, the paper would make an important contribution to research. Specific examples where editing would help clarify the paper's message: (Note: possible revisions or reviewer comments in italics; list not intended to be exhaustive). Abstract Except for all adult household members, we also explored the relationships among dyads of parents and children and spouses. Using a subsample of children with parents' chronic conditions as the key risk factor and a subsample of wives with the chronic conditions of the husband as key risk factor, we reran our GEE models to explore chronic condition concordance within these relationships (note: in any case drop reference to 'dyad'. It only helps to mislead the reader). This study provided the evidence about the effect of co-residence factor in the prevalence of chronic conditions. This study provides evidence that the chronic conditions of other members of a household may be a significant risk factor for a household member's own health (note: drop reference to 'co-residence'. The analyses do not include a 0/1 variable indicating co-residence) . Strength and Limitations

	This is the first study to estimate the risk of chronic conditions when family members has one or more chronic conditions in China. This is the first study in China to estimate the risk to a household member's own health, which associate with the chronic conditions of other household members. We performed a quantitative estimation of the assoication between one's chronic condition status and their household members, including parents-children and spouses. We perform multivariate logistic models to estimate the association between an adult's own chronic condition status and the chronic condition status of other household members. These models are run on several samples including all adult household members, children with parents' chronic conditions as key risk factor, and wives with husband's chronic conditions as key risk factor. Our study could not provide the estimation about the risk of new chronic condition in health household members and find out some significant risk factors. Based as it is on cross-sectional data, this study does not estimate the risk of a new chronic condition in a household member, nor does it provide evidence of a causal relationship. Introduction The life expectancy have increased, lifestyles have changed, health care have become more accessible, and health insurance coverage have increased. For these remarkable growths, the burden of disease for country and residents have shifted from some infectious diseases to non-communicable chronic conditions (2-4). Some researches had shown…. Life expectancy has increased, lifestyles have changed, health care has become more accessible, and health insurance coverage has increased. Aligned with these remarkable improvements, healthcare concerns in the country have expanded from a narrow focus on infectious diseases to encompass treatment for non-communicable chronic conditions as well (2-4). Research has shown…. Non-communicable chronic conditions were also main burden of diseases for disability-adjusted life-years (DALYs), high systolic blood pressure, high fasting plasma glucose, and high body mass index were the 1st, 2nd and 4 th ranked risk factor for globe DALYs in 2015, recepectively. While, in 1990, the ranks of these three diseases were 3rd, 10th and 13th, recepectively In measuring disability-adjusted life-years (DALYs), non-communicable chronic conditions have become main contributors to a country's burden of disease. As of 2015, high systolic blood pressure, high fasting plasma glucose, and high body mass index respectively ranked 1st, 2nd and 4th as important risk factors related to DALY. In 1990, the respective ranks of these three diseases were only 3rd, 10th and 13th. Additionally, detecting the risk factors of chronic conditions incidence and finding out the high-risk population are also the effective way to control the prevalence of chronic conditions. Additionally, detection of risk factors and identification of individuals at high-risk are important first steps in the prevention and treatment of chronic conditions. The chronic conditions associations would be detected among all household members, dyads of parents and children, and spouses, respectively. Risk factors based on chronic conditions in other family members were detected for samples consisting of all household members, children, and wives, respectively. (Reviewer note: Probably better

	to not make reference to 'dyads'. Dyads are single units and as such, you are not looking at associations among dyads, but rather associations between values within the dyad. Also, a reference to a dyad by itself does not provide necessary information about which household member is providing the outcome status and which is providing the value on the key factor. There are clearer ways to describe your analyses). Methods Five chronic conditions Hypertension:……. Diabetes:…… IHD: ….. CVD:…… Obesity:….. Reviewer comment: The 'Five chronic conditions' section just contains a glossary for the five conditions studied. The section should include some form of introduction and then a description of these five outcomes. Also, it could be shortened given the repetitiveness of how the chronic conditions were determined (most based on 'YES' survey responses). Covariates Some socio-demographic characteristics would be included in our analyses as covariates. Age (continuously specified in years), education status (illiteracy/primary, secondary or college), health insurance status (yes or no), smoking (yes or no), drinking (yes or no). Socio-demographic characteristics added to our models as covariates include age (continuously specified in years), education status (illiteracy/primary, secondary or college), health insurance status (yes or no), smoking (yes or no), and drinking (yes or no). Statistical analyses Before statistical analyses, we would identify whether each participant live in the household with the given chronic conditions. For example, if any other resident (excluding self) had the given chronic conditions, the indicator of household situation for he/she was "YES". Then the generalized estimating equations (GEE) model with logit link would be used to find out the relationship between one's chronic conditions and the others with chronic conditions living in the same household. Reviewer comment: The Statistical analysis section needs to be rewritten. The description of the analyses needs to: 1. specify the outcome (i.e., 0/1 variable for chronic condition status of the sampled subject) 2. specify the key risk factor studied (i.e., chronic condition status within group of other specified household member(s) – numerous groups are involved here such as males, females, parents, husbands, etc.) 3. specify the sample (e.g., all adult members, all children, all wives; repeatedly the paper notes what a subject needs to be valued as a '1', but does not devote space describing the full sample. In many cases it almost seems that only individuals with chronic conditions are used in the model). 4. specify the hierarchical structure of the model (model is described as GEE with logit link, but still need to know if it was two-level (individual within household), three level (individual within household within village), or four-level (individual within household within village within town). 5. specify if sampling made an effort to reduce inequality in probability of selection (for example, by using probability

	proportional to size) and whether analyses were weighted or not afterwards. 6. specify the working correlation matrix identified for use, not just that it was based on QIC (note: presumably matrix was exchangeable. If not, one might question appropriateness of QIC to determine). The results of unadjusted models would be obtained in supplementary information. Reviewer suggestion: Drop all reference to unadjusted models and the accompanying tables. Unadjusted models do not provide useful information except to provide the unscientific perception (often faulty) of the relationships being studied. Discussion Although we did not provide some specific risk factors for the prevalence of chronic conditions, the quantitative assessment of the associations had shown the effect of co-residence factor for the disease status among the household members. Reviewer comment: I do not understand this sentence and suggest you drop all reference to co-residence throughout the paper. In the first place, the authors do, in fact, provide specific risk factors for the prevalence of chronic conditions (namely, the presence of the chronic condition in another household member), and in the second place the authors do not show the effect of co-residence. To truly test the effect of co-residence you would randomly assign each sample subject a chronic condition patient and then test whether sample subjects, who were assigned co-residing patients through the randomization process, have a higher likelihood of having the chronic condition themselves. This would be a very different study. Patient and public involvement This study is a cross-sectional questionnaire survey and no patient involved. This is a cross-sectional study based on survey responses. It includes no further patient involvement.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Ian Fyffe

Institution and Country: Simon Fraser University, Canada

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

There is very little that I would suggest changing to this manuscript. One is changing the word "scene" to "sub-sample". Another is expanding the limitations section to include how household cohabitation in this instance includes health behaviour such as diet, exercise and second hand smoke as well as biological factors such as genetics. Although this was addressed earlier in the manuscript, I think it would be important to reiterate since the study is ultimately unable to pinpoint the underlying causes.

Reply:

We thank for the reviewer’s recognition.

For the first comment, we changed the word "scene" and “subset” to "subsample" in our article. We agreed that “subsample” would be more appropriate.

Thanks for the second comment. We added a sentence of “The mechanism and specific causes of cohabitation effects on prevalence of chronic conditions could not be presented by our study” in the third limitation:

“Third, our study could neither estimate the risk of new chronic conditions among household members nor find out a certain significant risk factor. We only pointed out the phenomenon that there were associations for the prevalence of chronic conditions at the household level. The mechanism and specific causes of cohabitation effects on prevalence of chronic conditions could not be presented by our study. These results suggested that we’d better pay attention to the health status of the health residents in the households with members with chronic conditions.”

Reviewer: 2

Reviewer Name: Danilo Silva

Institution and Country: Federal University of Sergipe - Brazil

Please state any competing interests or state ‘None declared’: None declared.

Please leave your comments for the authors below

Major points

Point 1: The study investigates an interesting issue regarding the co-occurrence of similar heath outcome in people living in the same household. Besides large sample size, my concern is whether sampling process was in accordance with the aim of the current study (more than one person per household). If not, I think some bias is possible. For instance, what differentiate houses that have more than one respondent?

Reply:

We thank for the reviewer’s comment. The sampling method of Health Service Survey of Shanghai was a three-stage, stratified, random sampling. In the third stage of sampling, a total of 12,002 households were included in the survey, with no regard of the number of respondents in one household. If one household was sampled, all members in this household were sampled. There is a question in the questionnaire, “How many person live in the house within 6 months”. We can identify the number of respondents in each house based on this question. In our study, the households with only one member were excluded, which might cause bias. For the households with one member, we could not obtain the information of chronic conditions for other person living in the same household. And the selection bias might not be completely avoided. Therefore, we added a sentence of “The exclusion of households with only one family member might also cause some selection bias” in the fourth limitation.

Point 2: The self-reported measures are understandable given the number of participants. However, mainly for chronic diseases, some people can have the disease but do not have the diagnostic. For

the current study, this could bring a perspective of the diagnosis and health care. When a member of the household receive the diagnostic of some chronic condition, are the other more prone to know and to perform diagnostic tests?

Reply:

Thank you for the comment. In our study, the diagnosis of chronic condition for participant was based on self-reported records, and the potential risk of underreporting and misreporting would be not avoided completely. We considered that this is the main limitation for our study. The rigorous design of study and formal training for investigators could partly reduce the impact of the bias. The reviewer mentioned that the participants are prone to know and to perform diagnostic tests, when a member of the same household receive the diagnostic of chronic conditions. We agreed that this phenomenon would be exist possibility. And this might induce reporting bias and also be one of possible reasons why the concordance of chronic conditions happen among the household members. We add discussion of “Besides, if a member of the household was diagnosed with a certain type of chronic condition, the likelihood of the other family members having their diagnostic tests performed was higher. This phenomenon might cause reporting bias and lead to the association of chronic conditions among household members” at the second paragraph of “Discussion” section. And the reporting bias had been mentioned in the fourth limitation: “Fourth, this study is a cross-sectional survey study, and several potential bias, such as recall bias, confounding bias and reporting bias, might not be completely avoided”.

Point 3: Are covariates actually moderators/mediators? I think the discussion should consider this. If people in the same household cluster diseases, this could be explained by some of the covariates or other not controlled such as diet and physical activity. Thus, I suggest study these exposures factors instead control them. This could also help the understanding about the discussion “environment vs genetic”.

Reply:

Thanks for the reviewer’s suggestion and we agreed with the comment. We added Table S5 “The results of full models for the association between chronic conditions of individual with the same condition of household member” in the supplementary materials to indicate the effect of these exposures. These were all the covariates for individuals. The results indicated that the effect of most of the covariates was consistent across three subsamples. We found positive association between age and incidence of chronic conditions except for obesity in subsample 2. And higher level of education status was associated with higher prevalence of any chronic conditions, hypertension, and diabetes. We added discussion of “In order to show the effects of covariates more clearly, we listed the results of full models for three subsamples in the Table S5. We found that the effects of most covariates were consistent across three subsamples” at the end of the third paragraph of “Discussion” section.

Minor point

Point 1: Information on China’s National Health Service Survey in the last paragraph of the introduction (page 5) should be referenced.

Reply:

We thank for your comment. We cite pervious research in the manuscript.

ref 17: Zhang X, Zhang Y, Xiao X, Ma X, He J. The relation between health insurance and management of hypertension in Shanghai, China: a cross-sectional study. BMC public health 2016;16:959 doi: 10.1186/s12889-016-3627-3.

The information of diabetes for the fifth Health Service Survey of Shanghai was calculated with this survey data. Therefore, there was no relative reference.

Reviewer: 3

Reviewer Name: Grant Ritter

Institution and Country:

Brandeis. University

Massachusetts, USA

Please state any competing interests or state 'None declared': None

Subject: Review of "Concordance of chronic conditions among the household members in Shanghai: a cross-sectional study"

Date: September 30, 2019

General Comments:

The research question is interesting and important, and the methods used in the analyses appear sound. The study is worth publishing. However, the current writing is not up to the usual level for a journal article and would benefit from further work. While many readers may be able to figure out much of the paper's content (particularly, if they study the accompanied tables and read the STROBE in the appendices), it needs another round of editing. Hopefully, the authors can be helped by someone with more experience in this area. Grammatical problems (e.g., verb tenses, subject-verb disagreement, dangling or misplaced modifiers, and sentence fragments) should be corrected, and the editor should also revise the content and sequence of written passages to provide better flow and improve readability.

Reply:

Thank you for the reviwer's positive comments to our manuscript. We appreciated that the reviewer had provided professional revision for our work, which have been very helpful in improving the manuscript. Additionally, we asked a native editing service called "Editage" to conduct a professional revision to our manuscript to make it more readable and fluent.

In particular, a full rewrite of the analysis plan - what the authors wish to test and how they plan to test it - would be beneficial. The paper uses a number of phrases to describe the analyses ('explored the relationships', 'estimate the risk of chronic conditions when family members has (sic) one or more chronic conditions', 'estimation of the association between one's chronic condition status and their household members' etc.), which are generally too vague. In the Methods section when the GEE logistical model is finally introduced, details are lacking regarding model specification and the

composition of the sample and subsamples used to test various hypotheses. With revision, the paper would make an important contribution to research.

Reply:

Thanks for the comment. We rewritten the statistical analysis plan and “Statistical analyses” section to make it provide more details and we add more information about the GEE model. We rephrased the description of analyses as your suggestion. And, we added “Objective” section to declare “The aim of our study is to estimate the association between an adult’s own chronic condition status and the chronic condition status of other household members” in SAP.

Specific examples where editing would help clarify the paper’s message:

(Note: possible revisions or reviewer comments in italics; list not intended to be exhaustive).

Abstract Except for all adult household members, we also explored the relationships among dyads of parents and children and spouses.

Using a subsample of children with parents’ chronic conditions as the key risk factor and a subsample of wives with the chronic conditions of the husband as key risk factor, we reran our GEE models to explore chronic condition concordance within these relationships (note: in any case drop reference to ‘dyad’. It only helps to mislead the reader).

Reply:

Thanks for the comment. We revised this part accordingly. We changed the expression of “dyads of parents and children” as “adult children subsample”.

This study provided the evidence about the effect of co-residence factor in the prevalence of chronic conditions.

This study provides evidence that the chronic conditions of other members of a household may be a significant risk factor for a household member’s own health (note: drop reference to ‘co-residence’. The analyses do not include a 0/1 variable indicating co-residence) .

Reply:

Thanks for the comment. We revised this part accordingly. We dropped all reference to “co-residence” in this paper.

Strength and Limitations

This is the first study to estimate the risk of chronic conditions when family members has one or more chronic conditions in China.

This is the first study in China to estimate the risk to a household member’s own health, which associate with the chronic conditions of other household members.

Reply:

Thanks for the comment. We revised this part accordingly.

We performed a quantitative estimation of the association between one's chronic condition status and their household members, including parents-children and spouses.

We perform multivariate logistic models to estimate the association between an adult's own chronic condition status and the chronic condition status of other household members. These models are run on several samples including all adult household members, children with parents' chronic conditions as key risk factor, and wives with husband's chronic conditions as key risk factor.

Reply:

Thanks for the comment. We revised this part accordingly.

Our study could not provide the estimation about the risk of new chronic condition in health household members and find out some significant risk factors.

Based as it is on cross-sectional data, this study does not estimate the risk of a new chronic condition in a household member, nor does it provide evidence of a causal relationship.

Reply:

Thanks for the comment. We revised this part accordingly.

Introduction

The life expectancy have increased, lifestyles have changed, health care have become more accessible, and health insurance coverage have increased. For these remarkable growths, the burden of disease for country and residents have shifted from some infectious diseases to non-communicable chronic conditions (2-4). Some researches had shown….

Life expectancy has increased, lifestyles have changed, health care has become more accessible, and health insurance coverage has increased. Aligned with these remarkable improvements, healthcare concerns in the country have expanded from a narrow focus on infectious diseases to encompass treatment for non-communicable chronic conditions as well (2-4). Research has shown….

Reply:

Thanks for the comment. We revised this part accordingly.

Non-communicable chronic conditions were also main burden of diseases for disability-adjusted lifeyears (DALYs), high systolic blood pressure, high fasting plasma glucose, and high body mass index were the 1st, 2nd and 4 th ranked risk factor for globe DALYs in 2015, respectively. While, in 1990, the ranks of these three diseases were 3rd, 10th and 13th, respectively。

In measuring disability-adjusted life-years (DALYs), non-communicable chronic conditions have become main contributors to a country's burden of disease. As of 2015, high systolic blood pressure, high fasting plasma glucose, and high body mass index respectively ranked 1st, 2nd and 4th as important risk factors related to DALY. In 1990, the respective ranks of these three diseases were only 3rd, 10th and 13th.

Reply:

Thanks for the comment. We revised this part accordingly.

Additionally, detecting the risk factors of chronic conditions incidence and finding out the high-risk population are also the effective way to control the prevalence of chronic conditions.

Additionally, detection of risk factors and identification of individuals at high-risk are important first steps in the prevention and treatment of chronic conditions.

Reply:

Thanks for the comment. We revised this part accordingly.

The chronic conditions associations would be detected among all household members, dyads of parents and children, and spouses, respectively.

Risk factors based on chronic conditions in other family members were detected for samples consisting of all household members, children, and wives, respectively. (Reviewer note: Probably better to not make reference to 'dyads'. Dyads are single units and as such, you are not looking at associations among dyads, but rather associations between values within the dyad. Also, a reference to a dyad by itself does not provide necessary information about which household member is providing the outcome status and which is providing the value on the key factor. There are clearer ways to describe your analyses).

Reply:

Thanks for the comment. We revised this part accordingly. We changed the expression of "dyads of parents and children" as "adult children subsample".

Methods

Five chronic conditions

Hypertension:…….

Diabetes:……

IHD: …..

CVD:……

Obesity:…..

Reviewer comment: The ‘Five chronic conditions’ section just contains a glossary for the five conditions studied. The section should include some form of introduction and then a description of these five outcomes. Also, it could be shortened given the repetitiveness of how the chronic conditions were determined (most based on ‘YES’ survey responses).

Reply:

Thanks for this benefit comment. We rephrased this section as following.

“Five chronic conditions

In this survey, we chose five chronic conditions with high prevalence rates: hypertension, diabetes, IHD, CVD, and obesity. The definition of these chronic conditions were based on the corresponding questions in the questionnaire, the disease coding list of the NHSS, and Body mass index (BMI). If a participant chose “YES” or a specific disease code, he or she was considered to have the corresponding chronic condition.

Hypertension: Hypertension was indicated based on the question “Have you ever been told by a doctor that you have hypertension?” in the questionnaire.

Diabetes: Diabetes was indicated based on the question “Have you ever been told by a doctor that you have diabetes?” in the questionnaire.

IHD: The disease codes for IHD included angina pectoris (061), myocardial infarction (062), and other ischemic heart disease (063).

CVD: The disease code for CVD included cerebrovascular disease (067).

Obesity: Obesity was indicated by the World Health Organization (WHO) International BMI categories (BMI ≥ 30 kg/m2)[21].”

Covariates

Some socio-demographic characteristics would be included in our analyses as covariates. Age (continuously specified in years), education status (illiteracy/primary, secondary or college), health insurance status (yes or no), smoking (yes or no), drinking (yes or no).

Socio-demographic characteristics added to our models as covariates include age (continuously specified in years), education status (illiteracy/primary, secondary or college), health insurance status (yes or no), smoking (yes or no), and drinking (yes or no).

Reply:

Thanks for the comment. We revised this part accordingly.

Statistical analyses

Before statistical analyses, we would identify whether each participant live in the household with the given chronic conditions. For example, if any other resident (excluding self) had the given chronic

conditions, the indicator of household situation for he/she was “YES”. Then the generalized estimating equations (GEE) model with logit link would be used to find out the relationship between one’s chronic conditions and the others with chronic conditions living in the same household.

Reviewer comment: The Statistical analysis section needs to be rewritten. The description of the

analyses needs to:

1. specify the outcome (i.e., 0/1 variable for chronic condition status of the sampled subject)

2. specify the key risk factor studied (i.e., chronic condition status within group of other specified household member(s) – numerous groups are involved here such as males, females, parents, husbands, etc.)

3. specify the sample (e.g., all adult members, all children, all wives; repeatedly the paper notes what a subject needs to be valued as a ‘1’, but does not devote space describing the full sample. In many cases it almost seems that only individuals with chronic conditions are used in the model).

4. specify the hierarchical structure of the model (model is described as GEE with logit link, but still need to know if it was two-level (individual within household), three level (individual within household within village), or four-level (individual within household within village within town).

5. specify if sampling made an effort to reduce inequality in probability of selection (for example, by using probability proportional to size) and whether analyses were weighted or not afterwards.

6. specify the working correlation matrix identified for use, not just that it was based on QIC (note: presumably matrix was exchangeable. If not, one might question appropriateness of QIC to determine).

Reply:

Thanks for the reviewer’s comment. We rephrased the “Statistical analyses” section as your suggestion. We described the primary outcomes and exposure (key risk factor) in the first paragraph. Then we presented the definition of the three subsamples and method for identifying status of exposure (“No” or “Yes”) in each subsample. We declared that two-level hierarchical structure of the model (individual within household) was considered in our study. Because there was no relevant information of sampling weight, the analyses were not weighted. For the working correlation matrix, we agreed with your suggestion and chose the exchangeable working correlation matrix for GEE model (“EXCH” option in SAS). Because the working correlation matrix for some analyses had been changed, the corresponding results were updated.

Statistical analyses

This study had five primary outcomes, each of which represented the status of each of the five chronic conditions, namely hypertension, diabetes, IHD, CVD, and obesity, in the participants (“No” or “Yes”). If a participant had any of these five chronic conditions (“Yes”), his or her status of having “Any chronic condition” is considered as “Yes”. To make a comprehensive assessment, we settled three subsamples: all adult household members (the total sample), adult children (adult children subsample), and wives (spouse subsample). In addition, we identified chronic conditions in the other household members as the exposure (or risk factor) in three subsamples.

For the first subsample of all adult household members, we included all household members aged 18 years or older. If any other residents (excluding self) have the given chronic conditions, the exposure status of household situation for each participant is identified as “Yes”. For the second subsample of adult children, only adult children would be included in analyses, and those participants were excluded if the disease information of parents were not available. The chronic conditions status of their parents were considered as exposure. For the third subsample of spouses, we included married women in the analyses. We defined the chronic conditions of wives as the outcomes and the chronic conditions of husbands as the exposure.

The generalized estimating equations (GEE) model with logit link would be used to explore the associations between chronic conditions of participants and the conditions of the others living in the same household. We considered a two-level hierarchical structure of the model (individuals within households). The model was based on individual’s data without taking sampling weight into account due to the lack of relevant information. The odds ratio (OR) and 95% confidence interval (CI) were estimated by the GEE model to indicate the association between any chronic condition or each given chronic condition of an individual and the same condition of household member (e.g., the association between the hypertensive status of a participants and that same condition in his or her other household members). Besides, the association of different chronic conditions would also be assessed (e.g., the association between hypertension in a participant and diabetes in other household members).

Adjusted models and unadjusted models were both used to estimate the associations in three household subsamples. The adjusted models included age, gender, health insurance status, education status, drinking and smoking; however, gender was excluded in the third subsample. The final conclusion was based on the results of the adjusted models. The results of unadjusted models would be obtained in supplementary information. We chose the exchangeable working correlation matrix for GEE models.

The subgroup analyses were conducted for the first subsample according to two pre-defined stratification factors: sex (male or female), education (illiteracy/primary, secondary, and college). We did not conducted any statistical model to deal with the missing data because of low missing data rates. Any observation with missing data would be excluded from the final analyses. All data management and statistical analyses were performed using SAS software (version 9.4; SAS Institute Inc., Cary, NC). All reported p values were two-sided and p value < 0.05 was regarded as statistically significant. This study was reported based on the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines (supplementary S1), and all analyses were conducted according to the statistical analysis plan (supplementary S2).

The results of unadjusted models would be obtained in supplementary information.

Reviewer suggestion: Drop all reference to unadjusted models and the accompanying tables. Unadjusted models do not provide useful information except to provide the unscientific perception (often faulty) of the relationships being studied.

Reply:

Thanks for the review’s comment, but we still keep the results of unadjusted models in our article. In the 16th item of the STROBE checklist, they required the unadjusted estimates: “(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval).” We agreed with you that unadjusted models do not provide the scientific and reliable

information. Therefore, we declared that the final conclusion was based on the results of the adjusted models, and the results of unadjusted models were only listed in the supplementary materials.

Discussion

Although we did not provide some specific risk factors for the prevalence of chronic conditions, the quantitative assessment of the associations had shown the effect of co-residence factor for the disease status among the household members.

Reviewer comment: I do not understand this sentence and suggest you drop all reference to co-residence throughout the paper. In the first place, the authors do, in fact, provide specific risk factors for the prevalence of chronic conditions (namely, the presence of the chronic condition in another household member), and in the second place the authors do not show the effect of co-residence. To truly test the effect of co-residence you would randomly assign each sample subject a chronic condition patient and then test whether sample subjects, who were assigned co-residing patients through the randomization process, have a higher likelihood of having the chronic condition themselves. This would be a very different study.

Reply:

Thanks for the reviewer's comment. We aimed to refer "co-residence factor" as a common living environment for those members living in the same household. However, the phrase of "the effect of co-residence factor" might cause misunderstanding that there was a variable as "co-residence factor" in our study. Therefore, we changed the sentence of "The main purpose of our study was to indicate the associations of chronic conditions among the household members. Although we did not provide some specific risk factors for the prevalence of chronic conditions, the quantitative assessment of the associations had shown the effect of co-residence factor for the disease status among the household members" into "*Although we did not provide certain specific risk factors for the prevalence of chronic conditions, the quantitative assessment showed the concordance of chronic conditions within households*".

We dropped all reference to "co-residence" in this paper.

Patient and public involvement

This study is a cross-sectional questionnaire survey and no patient involved.

This is a cross-sectional study based on survey responses. It includes no further patient involvement.

Reply:

Thanks for the comment. We revised this part accordingly.

VERSION 2 – REVIEW

REVIEWER	Danilo Silva Federa University of Sergipe – Brazil
REVIEW RETURNED	25-Nov-2019

GENERAL COMMENTS	The authors solved my concerns. Considering the limitations mentioned, I think this paper can contribute for the undestanding of chronic conditions.

REVIEWER	Grant A Ritter Brandeis University, Waltham, Ma. USA
REVIEW RETURNED	25-Nov-2019

GENERAL COMMENTS	Very responsive to previous comments. I include just a few small wording suggestions to clarify a few issues. Re: Review of 'Concordance of Chronic Conditions among Household Members….' Date: November 25, 2019 General comment: Paper authors were very responsive to review comments and paper is now acceptable. I think a few wording changes would help with small remaining issues in the paper and make it a bit clearer. Such comments, however, are minor, offered only as suggestions, and do not detract from the true value of the paper. Page 3 line 41-42: '……were included for study. Using all adult household members, we found……' Page 4 line 17 '……logistic GEE models …..' Page 8 line 56 '…….we employed three models……' The identification of the various samples used for models could be confusing. The authors refer to these samples at various times as the 'first subsample', 'second subsample' and 'third subsample'. This could make a reader query what constituted the full sample, and where was its corresponding analysis and results. It would be better to call the sample of all adult household members the 'full study sample' and only identify two subsamples (adult children and spouse). This would need to be done throughout the paper beginning on page 9: Page 9 line 8 - '……using the full study sample of all adult household members, we…..' Page 9 line 15-16 ' …For the first subsample of adult children…….For the second subsample of spouses…...' Etc. Continuing with other suggestions: Page 9 line 28 '…….generalized estimating equation model with ……' Page 9 line 47 '…….household members). In addition, the association of…' Page 10 line 6 '…..results of unadjusted models are given to provide supplemental information. Page 10 line 8 '….correlation matrix to estimate standard errors of coefficients in the GEE models.' Page 10 line 52 ' …..For the subsample of adult children, there were 5,489 available records, and for the subsample of spouses, there were 7,844 records.' Page 13 line 30 '….disease. .However, if a member of the household …' Page 13 line 56 to page 14 line 9 '…. to further explore the effect of a common genetic factor or a common living environment, we analyzed two subsamples, a subsample of adult children to exam the two factors together, and a subsample of spouses to exam living environment factors in the absence of a common genetic

	link. Similar to results using all adulthousehold members, positive associations were found …. ‘ Page 15 lines 10-11 - I’m not sure what you mean with this sentence since certain specific risk factors were used in your models. Perhaps you are thinking of omitted risk factors. I recommend that you rewrite this sentence to make it more clear,

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Danilo Silva
Institution and Country: Federa University of Sergipe - Brazil
Please state any competing interests or state ‘None declared’: None declared.

Please leave your comments for the authors below
The authors solved my concerns. Considering the limitations mentioned, I think this paper can contribute for the undestanding of chronic conditions.
Reply:
Thanks for the reviewer’s recognition for our work.

Reviewer: 3
Reviewer Name: Grant A Ritter
Institution and Country: Brandeis University, Waltham, Ma. USA
Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
Very responsive to previous comments. I include just a few small wording suggestions to clarify a few issues.
General comment: Paper authors were very responsive to review comments and paper is now acceptable. I think a few wording changes would help with small remaining issues in the paper and make it a bit clearer. Such comments, however, are minor, offered only as suggestions, and do not detract from the true value of the paper.
Reply:
Thanks for the reviewer’s recognition for our work. We appreciated that the reviewer provided assistance of the expression improvement for our study.

Page 3 line 41-42: ‘……were included for study. Using all adult household members, we found……’
Reply:
Thanks for the comment. We revised this part accordingly.

Page 4 line 17 ‘……logistic GEE models …..’
Reply:
Thanks for the comment. We revised this part accordingly.

Page 8 line 56 ‘…….we employed three models……’
Reply:
Thanks for the comment. We thought that three subsamples would be appropriate and “three models” might be misunderstood as three different statistical methods.

The identification of the various samples used for models could be confusing. The authors refer to these samples at various times as the ‘first subsample’, ‘second subsample’ and ‘third subsample’.

This could make a reader query what constituted the full sample, and where was its corresponding analysis and results. It would be better to call the sample of all adult household members the 'full study sample' and only identify two subsamples (adult children and spouse). This would need to be done throughout the paper beginning on page 9:
Page 9 line 8 - '……using the full study sample of all adult household members, we…..'
Page 9 line 15-16 ' …For the first subsample of adult children…….For the second subsample of spouses…...'
Reply:
Thanks for the comment. We had changed "the first subsample" into "the full study sample of all adult household members", "second subsample" into "the subsample of adult children", and "third subsample" into "the subsample of spouses". The revisions were shown as below:
"For the full study sample of all adult household members, we included all household members aged 18 years or older." (the second paragraph of "Statistical analyses" section)
"For the subsample of adult children, only adult children..." (the second paragraph of "Statistical analyses" section)
"For the subsample of spouses, we included married women in the analyses" (the second paragraph of "Statistical analyses" section)
"…gender was excluded in the subsample of spouses." (the fourth paragraph of "Statistical analyses" section)
"The subgroup analyses were conducted for the full study sample of all adult household members according to two pre-defined stratification factors…" (the last paragraph of "Statistical analyses" section)
"However, the adult children participants in the subsample of adult children were young (36.01 years)…" (the first paragraph of "Results" section)
"In the full study sample of all adult household members, the results indicated that the chronic conditions…" ("All household members" part of "Results" section)
"The results of unadjusted GEE models for the subsample of adult children were listed in the Table S3." ("Adult children subsample" part of "Results" section)

Continuing with other suggestions:
Page 9 line 28 '…….generalized estimating equation model with ……'
Reply:
Thanks for the comment. It is the first time this abbreviation (GEE) has appeared in the manuscript. So we considered that both of the full expression and abbreviation need to be preserved.

Page 9 line 47 '…….household members). In addition, the association of…'
Reply:
Thanks for the comment. We revised this part accordingly.

Page 10 line 6 '…..results of unadjusted models are given to provide supplemental information.
Reply:
Thanks for the comment. We revised this part accordingly.

Page 10 line 8 '….correlation matrix to estimate standard errors of coefficients in the GEE models.'
Reply:
Thanks for the comment. We revised this part accordingly.

Page 10 line 52 ' …..For the subsample of adult children, there were 5,489 available records, and for the subsample of spouses, there were 7,844 records.'
Reply:
Thanks for the comment. We revised this part accordingly.

Page 13 line 30 '….disease. .However, if a member of the household …'
Reply:
Thanks for the comment. We revised this part accordingly.

Page 13 line 56 to page 14 line 9 '…. to further explore the effect of a common genetic factor or a common living environment, we analyzed two subsamples, a subsample of adult children to exam the two factors together, and a subsample of spouses to exam living environment factors in the absence of a common genetic link. Similar to results using all adult household members, positive associations were found …. '
Reply:
Thanks for the comment. We revised this part accordingly.

Page 15 lines 10-11 - I'm not sure what you mean with this sentence since certain specific risk factors were used in your models. Perhaps you are thinking of omitted risk factors. I recommend that you rewrite this sentence to make it more clear.
Reply:
Thanks for the comment. The aim of our study is not to investigate certain specific risk factors for the prevalence of chronic conditions. We included the covariates in our analysis model. The expression of "we did not provide certain specific risk factors" might cause misunderstanding. Therefore, we rephrased this sentence into "Although the aim of our study is not to investigate certain specific risk factors for the prevalence of chronic conditions, the quantitative assessment showed the concordance of chronic conditions within households".