PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association of socioeconomic status with prognosis in hypertensive patients over age 65: A cohort study in the community setting
AUTHORS	Martin-Fernandez, Jesus; Alonso-Safont, Tamara; Gestri, Patricia Elena Mora; Polentinos Castro, Elena; Rodríguez-Martínez, Gemma; Bilbao, Amaia; Cura-Gonzalez, Isabel

VERSION 1 - REVIEW

REVIEWER NAME	Liu, Jue		
REVIEWER AFFILIATION	Peking University School of Public Health, Department of		
	Epidemiology & Biostatistics		
REVIEWER CONFLICT OF	None		
INTEREST			
DATE REVIEW RETURNED	18-Jul-2023		

GENERAL COMMENTS	In this manuscript, authors discussed the interesting topic about the association of socioeconomic status with prognosis of patients aged over 65 with hypertension. Though the relationships between socioeconomic status and hypertension or cardiovascular diseases have been studied widely, the effects of socioeconomic factors on outcomes of hypertensive patients are still unclear. As such, this paper is a welcome addition to the field.
	Major comments: 1. In this paper, the socioeconomic status was measured by social deprivation index. In my view, it would be better to add some descriptions of social deprivation index or MEDEA project or area- level socioeconomic status in the Introduction section. 2. Authors chose kidney diseases and cardiovascular diseases as primary outcomes to indicate the prognosis of hypertension. I think authors need supply reasons for this choice in the Introduction section, for example, the high incidence of cardiovascular diseases or kidney disease or the adverse outcomes of these disease among hypertensive patients. 3. The structure of Methods section need be modified, especially the presentation of variables. It would be easily understood if authors describe collected variables by groups as exposure, outcomes, and covariates. Additionally, it seemed that authors identified the primary outcomes as the incidence of kidney diseases and cardiovascular diseases and the mortality caused by these diseases. It might be possible that patients had records of both kidney diseases and cardiovascular diseases, or both kidney/cardiovascular diseases and related death, then which outcome is the first one concerned in this study? The discrimination of outcomes should be described more clearly.
	4. Some contents in the Discussion section should be transferred in the Introduction section, for example, the paragraph beginning with "One element to assess is the measurement of the socioeconomic

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

 situation". 5. The Discussion section is a part mainly including comparation of similar studies, interpretation of potential mechanisms about the association and additional values for policy based on the results in this study. It could not present any results which cannot be obtained from this study. For example, in the second paragraph, it would be better if authors simplified the similar results from other studies and added discussion about the mechanisms about the association between social deprivation and risk of kidney or cardiovascular diseases. In addition, "the association between socioeconomic status and mortality when HTN is combined with another cardiovascular risk factors" cannot be observed in this study, so it was improper to mention "our results also go in the same direction". If authors planned to explore the effects of DM on association between socioeconomic status and adverse outcomes of hypertension, subgroup analyses could be used. 6. Though the paragraph before Conclusion section presented the implication of these results on hypertension management, it was general in any other studies among patients with hypertension. Authors should highlight the specific constructive suggestion for policy from results in this study, such as the favorable interventions among vulnerable population (the older people, patients with low socioeconomic status).
 Minor comments: 1. Some citations of other studies lacked references. For example, "In studies carried out in our environment, it has been observed that the differences in the prevalence of hypertension according to socioeconomic factors in an older population are small" in Introduction, "In general, the trend of the past two decades is that inequalities in total mortality are reduced in all European countries and, especially, in Spain" in Discussion etc. 2. Authors need add the reasons for analysis by sex. 3. In Results, "The prevalence of smokers is lower than in the general population" cannot be obtained by data used in this study. 4. It was not clear the measuring unit of hypercholesterolemia in Table 1. Is that counted by number of patients? If is, the "." should be changed to ",". 5. The "," of HR and its confidence interval in Table 2 should be ".". 6. In abstract, the Intervention part can be deleted as this study was a cohort design.

REVIEWER NAME	Peng, Ke
REVIEWER AFFILIATION	Fuwai Hospital Chinese Academy of Medical Sciences
REVIEWER CONFLICT OF	I declare there is no competing interests.
INTEREST	
DATE REVIEW RETURNED	29-Aug-2023

GENERAL COMMENTS	This is an important paper contributing to the literature. The authors explored the associations between SES and incidence of kdney or cardiovascular events and all-cause mortality. However, I have a few comments on this study:
	Major: Page 9 line 10. MEDEA project used an index consisted of work and

education information. My concern is whether this index is validated to be used to indicate SES? As you mention this variable is related to general mortality, but what your research interest is mainly disease burden.
Page 9 line 50-52, the authors stated that 'Additionally, other concerns such as influential outliers, missing data, or significant model misspecification were considered.'. How they deal with the data, what are the percentages of each exception values were not presented, whether the yielded data was still balanced from the original one? How you deal with this data?
The authors should ask for a proof-reading from an epidemiology researcher who is an English native speaker.
Minor: Abstract:
MEDEA spell it out.
Use the terminology consistently through your paper, as you go socioeconomic statsus in objective, then use it as well in your results.
Introduction Should add some information of your SES index in this section.
Should add your research aim in your last paragraph.
Methods Indicate what software you use in the analysis section.
You may need to list a bit more on the inclu/exclu criteria.
Results: 12,334+9,419=21753 not 21754
We always use mean+SD to present age, but why you present IQR as well? Does age have a skewed distribution?
Have you done the difference test for table 1? Can you show the P value?
Table 2, should be 95% CI, p should be italicized., uppercase n of number of cluster. Suggest to correct it through your manuscript.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Prof. Jue Liu, Peking University School of Public Health

Comments to the Author:

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

In this manuscript, authors discussed the interesting topic about the association of socioeconomic status with prognosis of patients aged over 65 with hypertension. Though the relationships between socioeconomic status and hypertension or cardiovascular diseases have been studied widely, the effects of socioeconomic factors on outcomes of hypertensive patients are still unclear. As such, this paper is a welcome addition to the field.

Thanks for your considerations

Major comments:

 In this paper, the socioeconomic status was measured by social deprivation index. In my view, it would be better to add some descriptions of social deprivation index or MEDEA project or area-level socioeconomic status in the Introduction section.

MEDEA index construction has been described in *Domínguez-Berjón FM et al. Constructing a deprivation index based on census data in large Spanish cities [the MEDEA project] Gac Sanit.* 2008;22(3):179-8. Several socio-economic indicators were defined and their dimensionality studied. The indicators with the strongest correlations with overall mortality were those related to work, education, housing conditions and single parent homes.

In the analysis of dimensionality, a first dimension appeared that contained indicators related to work (unemployment, manual and eventual workers) and education (insufficient education overall and in young people). In all the cities studied, the index created with these 5 indicators explained more than 75% of their variability. The correlations between this index and mortality generally showed higher values than those obtained with each indicator separately.

We have added this paragraph in the Introduction section:

The study of socioeconomic status can be approached from an individual or contextual perspective. There are multiple characteristics that can define the socioeconomic situation, which can be considered from a multidimensional perspective(28). Many of the studies mentioned use indicators of individual socioeconomic status(23,25,29). Others use the socioeconomic status of the area and evaluate its relationship with survival after a cardiovascular event(30). In fact, the deprivation of the area has been shown to be a better predictor than the individual socioeconomic situation when studying the occurrence of cardiovascular events(31). In our setting, the MEDEA project ("Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconomic level, which has been shown to explain differences in mortality(32). Addressing socioeconomic differences from a contextual perspective has been successful in investigating differences in cardiovascular mortality due to certain diseases(30).

This change is intended to respond to your suggestions in this point and in point 4.

2. Authors chose kidney diseases and cardiovascular diseases as primary outcomes to indicate the prognosis of hypertension. I think authors need supply reasons for this choice in the Introduction section, for example, the high incidence of cardiovascular diseases or kidney disease or the adverse outcomes of these disease among hypertensive patients.

Thank you for your comment. Hypertension is a well-known risk factor for progression of kidney disease and cardiovascular events, including stroke, heart failure and myocardial infarction. Indeed, most of the burden of disease associated with hypertension is mediated by the circumstances

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

outlined above. We agree with the reviewer that this idea should be mentioned in the manuscript. We have added the following sentence in the introduction section:

Most of this burden of disease is due to cardiovascular and renal complications and mortality from these causes associated with the HTN diagnosis (14).

3. The structure of Methods section need be modified, especially the presentation of variables. It would be easily understood if authors describe collected variables by groups as exposure, outcomes, and covariates. Additionally, it seemed that authors identified the primary outcomes as the incidence of kidney diseases and cardiovascular diseases and the mortality caused by these diseases. It might be possible that patients had records of both kidney diseases and cardiovascular diseases, or both kidney/cardiovascular diseases and related death, then which outcome is the first one concerned in this study? The discrimination of outcomes should be described more clearly.

We have reformulated the methods section, changing the presentation of the study variables, grouping them together, and we have specified which event was studied as the main outcome in each model.

The variables included have been classified into exposure variable, covariates and outcome variables, as you suggested.

4. Some contents in the Discussion section should be transferred in the Introduction section, for example, the paragraph beginning with "One element to assess is the measurement of the socioeconomic situation ...".

We have made this change, see point 1, thanks.

5. The Discussion section is a part mainly including comparation of similar studies, interpretation of potential mechanisms about the association and additional values for policy based on the results in this study. It could not present any results which cannot be obtained from this study. For example, in the second paragraph, it would be better if authors simplified the similar results from other studies and added discussion about the mechanisms about the association between social deprivation and risk of kidney or cardiovascular diseases. In addition, "the association between socioeconomic status and mortality when HTN is combined with another cardiovascular risk factors..." cannot be observed in this study, so it was improper to mention "our results also go in the same direction". If authors planned to explore the effects of DM on association between socioeconomic status and adverse outcomes of hypertension, subgroup analyses could be used.

We have partially redrafted this paragraph. We have included some of the suggested mechanisms by which social inequality is associated with a worse prognosis of hypertension.

This is the sentence we added:

Suggested mechanisms to explain the association between socioeconomic disadvantaged environments and cardiovascular disease relate to dietary habits, physical activity resources and other cardiovascular risk factors(28). In this paper we evaluated the association between area-level socioeconomic status and kidney and cardiovascular events in hypertensive patients, adjusting for the effect of other risk factors such as diabetes, smoking and hypercholesterolemia. In this paper we evaluated the association between area-level socioeconomic status and kidney and cardiovascular events in hypertensive patients, adjusting for the effect of other risk factors such as diabetes, smoking and hypercholesterolemia.

On the other hand, we agree with your statement, we have not studied the combined effect of different risk factors with low socioeconomic status on the occurrence of events in patients with HTN so we have removed that sentence from the discussion.

6. Though the paragraph before Conclusion section presented the implication of these results on hypertension management, it was general in any other studies among patients with hypertension. Authors should highlight the specific constructive suggestion for policy from results in this study, such as the favorable interventions among vulnerable population (the older people, patients with low socioeconomic status).

We have rewritten the paragraph setting out the implications of the study, including your suggestion:

The implications of the results discussed relate to practice, health policies and research. Firstly, older patients diagnosed with hypertension in socioeconomic disadvantaged settings should be monitored particularly closely. Secondly, as has been suggested, to reduce the burden of disease derived from HTN, strategies are needed to accelerate the socioeconomic improvements of the most vulnerable population and the development of environments that promote health (47). Public health policies should focus on reducing social inequalities as a mechanism for improving individual health, with special attention to elderly patients. Finally, these results encourage further study of the role of social support, the cultural context of care and the health care system in the prognosis of these diseases.

The conclusion has also been redrafted:

Living in a low socioeconomic status area is associated with an increase in kidney or CV events in hypertensive patients diagnosed after the age of 65 years and with no previous cardiovascular history, which will result in a significant increase in disease burden even if not related to an increase in total mortality.

Minor comments:

1. Some citations of other studies lacked references. For example, "In studies carried out in our environment, it has been observed that the differences in the prevalence of hypertension according to socioeconomic factors in an older population are small..." in Introduction, "In general, the trend of the past two decades is that inequalities in total mortality are reduced in all European countries and, especially, in Spain" in Discussion etc.

The reference of the first sentence was placed at the end of the paragraph (reference 22). We have included two ideas from the same source and referenced them at the end.

The same is true in the second case. Reference 43 supports what has been written in the two previous sentences.

2. Authors need add the reasons for analysis by sex.

We have mentioned these reasons, adding a reference in Methods section

Gender differences play an influential role in multiple health-related outcomes. The importance of studying gender in investigating the role of the cultural and social environment in the prognosis of HTN has been highlighted(37).

3. In Results, "The prevalence of smokers is lower than in the general population..." cannot be obtained by data used in this study.

The sentence has been deleted.

4. It was not clear the measuring unit of hypercholesterolemia in Table 1. Is that counted by number of patients? If is, the "." should be changed to ",".

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

You are right, the number stand for diagnoses of hypercholesterolemia. The change has been made.

5. The "," of HR and its confidence interval in Table 2 should be ".".

Thank you, it was a mistake, the change has been done.

6. In abstract, the Intervention part can be deleted as this study was a cohort design.

Done

Reviewer: 2

Dr. Ke Peng, Fuwai Hospital Chinese Academy of Medical Sciences

Comments to the Author:

This is an important paper contributing to the literature. The authors explored the associations between SES and incidence of kidney or cardiovascular events and all-cause mortality. However, I have a few comments on this study:

Thank you for your appreciation.

<u>Major:</u>

Page 9 line 10, MEDEA project used an index consisted of work and education information. My concern is whether this index is validated to be used to indicate SES? As you mention this variable is related to general mortality, but what your research interest is mainly disease burden.

Thank you for your comment.

The MEDEA index has been used as a proxy indicator of area SES in multiple settings and studies (Bennett M et al. BMJ Open 2023 <u>http://dx.doi.org/10.1136/bmjopen-2022-066404</u>; Sáenz-Herrero M et al. Arch Womens Ment Health 2023 https://doi.org/10.1007/s00737-023-01360-x; Buron A, et al . Plos One 2017 <u>https://doi.org/10.1371/journal.pone.0179864</u>).

Its association with variability in mortality is well established (Nolasco A et al. Int J Equity Health 2015https://doi.org/10.1186/s12939-015-0164-0; Domínguez Berjon M et al Gac Sanit 2008 http://scielo.isciii.es/scielo.php?).

Additionally, their relationship with the variability of disease burden has been studied in the case of osteoarticular diseases or heart failure, for example. (Reyes C et al. Bone 2015 <u>https://doi.org/10.1016/j.bone.2014.12.019</u>; Garcia R, et al. J Epidemiol Community Health 2018 <u>http://dx.doi.org/10.1136/jech-2017-210146</u>).

Thus, we consider this variable to be useful both for studying differences in mortality and variability in burden disease.

Page 9 line 50-52, the authors stated that 'Additionally, other concerns such as influential outliers, missing data, or significant model misspecification were considered.'. How they deal with the data, what are the percentages of each exception values were not presented, whether the yielded data was still balanced from the original one? How you deal with this data?

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Extreme survival data were reviewed to ensure the veracity of the information prior to starting the analysis. The nature of the information sources used did not suggest a lack of information on outcome variables. If there was no information on the covariates, they were considered not to be present. If there was no record of the patient smoking in the clinical history, they were considered non-smokers, if there was no record of the patient being diabetic, they were considered non-diabetic, and so on with all the adjustment variables. All the data of the subjects included have been presented, (see figure 1), and the database used to carry out the analysis is provided so that these cases can be evaluated with complete transparency.

We assessed whether the specifications of the model could be affected by the proportionality assumptions. Since the Cox model requires the risks of the groups under study to be proportional, it should be tested whether the risks are proportional. As it is known, this can be done by graphical methods, or by studying the residuals. The latter system uses significance tests as a decision rule, which increase the probability of being significant, and therefore of discouraging the use of the method, the larger the sample size.

In a real-world setting, it is difficult to ensure that risks are perfectly proportional throughout the followup. In practice, this is not the case in most clinical situations.

We first study the residuals beyond significance testing.

Regarding the association between socioeconomic index and cardiovascular events, including cardiovascular mortality (Table 2), we found a violation of the proportional hazard assumption only for the hazard of 3rd quintile vs. 1st quintile (p<0.0001). For the rest, the proportional hazard assumption is met (p>0.05). We explored graphically the Schoenfeld residuals for all hazards (q2 vs q1, q3 vs q1, q4 vs q1 and q5 vs q1). As we can see in the following Figure, the residuals are pooled randomly on either side of the Y axis, and the fitted curve is close to a straight line.



The same graphics are plotted limiting the Y axis to -1 and 1, in the following figure. Although the downward trend is slightly greater for the hazard of 3rd quintile vs 1st quintile than for the rest, it is not very pronounced, with non-substantial violations of hazard proportionality.



Afterwards, following the recommendations of some authors (Stensrud MJ et al. JAMA 2020 doi: 10.1001/jama.2020.1267), we made an assessment of the influential outliers and potential model misspecification by testing the proportionality of risk in different brackets, prior to the use of the proposed method.

We show the HR at different time points, t≤month 48, t≤month 96 and t<end point, after splitting observation times.

The results studying the association between the MEDEA index and cardiovascular events, including cardiovascular mortality, are described below (table 2). It can be observed that HRs remain fairly stable, especially if we look at the least advantaged groups (fourth and fifth quintile). For the third quintile, the HR decreases, but only slightly, from 1.275 (95% CI, 1.044 – 1.556) to 1.148 (95% CI, 1.036 – 1.273).

HRs maintained direction and there were only slight-moderate variations over time, so we assumed the principle of proportionality, although the residual analysis significance tests showed some alerts.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Variable	HR	HR CI 95%	p> z
t<= month 48			
Socioeconomic group			<0.001
2nd vs. 1st quintile	1.011	0.855-1.196	0.855
3rd vs. 1st quintile	1.275	1.044-1.556	0.017
4th vs. 1st quintile	1.148	1.067-1.262	0.140
5th vs. 1st quintile	1.234	1.089-1.302	0.008
t<= month 96			
Socioeconomic group			<0.001
2nd vs. 1st quintile	1.014	0.912-1.128	0.798
3rd vs. 1st quintile	1.198	1.057-1.357	0.005
4th vs. 1st quintile	1.195	1.080-1.323	0.001
5th vs. 1st quintile	1.233	1.111-1.369	<0.001
t< final			<0.001
Socioeconomic group			
2nd vs. 1st quintile	1.009	0.917-1.111	0.847
3rd vs. 1st quintile	1.148	1.036-1.273	0.009
4th vs. 1st quintile	1.160	1.067-1.262	0.001
5th vs. 1st quintile	1.191	1.089-1.302	<0.001

Besides, we followed the recommendation to calculate 95% confidence intervals using bootstrapping techniques to improve the robustness of the models (Stensrud MJ et al. JAMA 2020 doi: 10.1001/jama.2020.1267).

The authors should ask for a proof-reading from an epidemiology researcher who is an English native speaker.

We have corrected some of the wording (all marked up) for a better understanding of the manuscript. However, since English is not our mother language, we entrusted the translation to AJE, who gave us a certificate of suitability for the assessed version (verification code ED2F-794A-327C-F046-9D7P).

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

Minor:

Abstract:

MEDEA spell it out.

It has been done

Use the terminology consistently through your paper, as you go socioeconomic status in objective, then use it as well in your results.

Done

Introduction

Should add some information of your SES index in this section.

We did it at the suggestion of both reviewers.

Should add your research aim in your last paragraph.

We have reworded the last paragraph in the Introduction section:

Therefore, we aimed to evaluate the potential association between the area-level socioeconomic status and the risk of kidney and cardiovascular events and mortality after the diagnosis of HTN, in a population aged 65 and older in the community setting.

Methods

Indicate what software you use in the analysis section.

The following sentence has been added at the end of methods section:

Stata 14.2 ® software was used for data analysis

You may need to list a bit more on the inclu/exclu criteria.

We have added exclusion criteria

The exclusion criteria were being under 65 years of age, suffering from kidney or cardiovascular disease, having been diagnosed with hypertension or taking antihypertensive medication before the start of the observation period.

Results:

12,334+9,419=21753 not 21754

Thanks, it was a mistake. In 2007, 12,335 patients with a diagnosis of HTN were included. It has been modified

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

We always use mean+SD to present age, but why you present IQR as well? Does age have a skewed distribution?

We show below the histogram for age. Since the distribution is amputated below 65 years of age, we believe that it is better characterized by the median and the interquartile range. In any case, the mean age and its SD are reported in the text.



Have you done the difference test for table 1? Can you show the P value?

We're sorry, but we're not sure it's appropriate to show inferential statistics in table 1, since the STROBE statement discourages this practice:

"Inferential measures such as standard errors and confidence intervals should not be used to describe the variability of characteristics, and significance tests should be avoided in descriptive tables. Also, P values are not an appropriate criterion for selecting which confounders to adjust for in analysis; even small differences in a confounder that has a strong effect on the outcome can be important" (Vandenbroucke JP et al. Plos Medicine 2007 *doi:10.1371/journal.pmed.0040297.t001*).

Table 2, should be 95% CI, p should be italicized., uppercase n of number of cluster. Suggest to correct it through your manuscript.

Done, thank you.

During the editorial process of this manuscript, new evidence on the validity of acute myocardial infarction and stroke diagnoses in our primary care medical record was published (De Burgos-Lunar et al, BMJ2023, doi: 10.1136/bmjopen-2022-068938), so the "Strengths and limitations of this study" section has been modified.

We welcome all suggestions, as we hope they will improve the manuscript.

On behalf of the authors

Dr. Jesús Martín-Fernández