PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prevalence and associated risk factors of hypertension among persons aged 15-49 in India: a cross sectional study
AUTHORS	Ghosh, Soumitra; Kumar, Manish

VERSION 1 – REVIEW

REVIEWER	Manuel Serrano Ríos MD.PhD. Emeritus Professor Medicine. Madrid, Spain. Complutense University- Madrid. Spain
REVIEW RETURNED	27-Mar-2019

GENERAL COMMENTS	The paper by titled “Prevalence and associated risk factors of hypertension in India: Evidence from NFHS” was aimed to review the global and interstates prevalences of Hypertension in India” by Ghost, Souintra. This a crossection study on data from the 4th round of the National Family Health Survey (NFHS) in India. “Data ……the only source of data that provides estimates on the social and demographic indicators up to the district level. It is a multistage stratified random sample survey, which gathers data primarily on demographic, socioeconomic and reproductive and child health (RCH) parameters but in the latest round, the scope has been widened with the inclusion of clinical, anthropometric and bio-chemical (CAB) tests and measurements of blood glucose and blood pressure (BP) for assessing the prevalence of non-communicable diseases such as diabetes and hypertension in the population. Additionally, it also collected information on behavioural risk factors such as consumption of alcohol and smoking. NFHS 4 was conducted during the period between 2015 and 2016 across all 29 states and 7 Union Territories (UTs) in India. …… The definition was based on the criteria given by WHO and American Heart Association (Pickering et al 2005). To make the prevalence of hypertension comparable, age adjusted prevalence rates were calculated for all states, UTs and districts. Apart from calculating the prevalence of hypertension, multivariate logistic regression model was employed to assess the correlates of hypertension.” The issue is indeed, interesting. Comments The currents draft is too long requiring be shorten in some specific sections (e.g: Methods could be reduced, Discussion). Also, number of the tables (4) and figures will likely be recommended sometime legends are scarcely explained or confusing. Also the English language requires revision by an expert.

	The main findings of this study should also (e.g. conclusions) neatly transcribed as: points a)…b)…c)…..d). In the current draft these key factors are, only maximally mentioned. To place tables and figure at the end (not in) as a supplement, of the written test. We would recommended to delete. Tables: a) Divide Table 1 (the real one in page 7) into 2 parts. (Education as the dividing point). And the now referred as Table 1, really Table 2 (page 8). b) To delete Tables A1 which is too complicated and whose content may essentially described in the text. c) Figures 1,2,4,5 could be suppressed or as complementary supplement. We recommended that the selected figures and tables be placed at the end of each page of the manuscript rather than inside the writing text at it appears now. We gratefully acknowledge the fact that the strength and limitations /some of them) of this study are recognized explicitly in page 3 (upper part). Limitations of the study ¬ The use of cross-sectional data that does not allow for exploration of causal pathways underlying the reported associations ¬ The role of behavioural risk factors such as low fruit and vegetable intake and physical inactivity could not be explored in this analysis ¬ Findings are limited to the adults aged between 15 and 54. Final comment on the contribution of the prevalence of Hypertension to Health in the different population of India. Possible ethnicity and genetic variation between people rural vs urban setting should be better described and better understanding the reported differences found in the prevalence of Hypertension between Northam and Southam states including their respective districts.

REVIEWER	Sarah-Jo Sinnott London School of Hygiene and Tropical Medicine, UK
REVIEW RETURNED	02-Apr-2019

GENERAL COMMENTS	General Overall, this is an important piece of research that estimates prevalence of hypertension in India at various geographical levels. The authors correctly state that having access to prevalence estimates is important for managing and measuring the effects of interventions to try help focus efforts on lowering prevalence rates. However, this paper needs substantial revision before it can be published. In brief, it needs much more details on the methods used, in particular how the survey was set up, administered, how the population was recruited, why the age ranges were selected for different genders, why there is a large proportion of young people in the sample. The large proportion of young people is of particular concern as it means the prevalence estimates generated are unlikely to be representative of the general population – indeed at 10% prevalence this is clearly reinforced. Global burden of disease estimates for hypertension are much higher than this. Even with clarity on how the population was set up the authors will need to be very clear on how these estimates generalise to the Indian population. The authors may want to consider weighting to make the estimates more generalisable. On a technical note there are some typos and some revisions of English may be required.

	Abstract The conclusions are not based on the data presented in the results section i.e. a statement is made about prevalence amongst poor people but there is no explanation for how “poor” is defined in the data or any results presented for “poor” vs “non-poor”. Introduction There are a few acronyms that have not been defined. Please make sure to spell out in first instance of use. A statement is made referring to inter-state variation in hypertension prevalence, based on differences in deprivation/socio-economic status but no reference is given for either SES or hypertension. Given that the authors also state that this study is the first examination of interstate prevalence I think some clarity here would be helpful; I expect the authors assume differences in hypertension prevalence based on differences in SES etc. Methods Much more detail is needed in the methods. Regarding the survey, how was the sample derived? How was the survey administered? Were people visited in their homes by trained survey teams? Were teams trained in anthropometric measurements? Was the survey administered via interview or self-complete? How does the respondent rate match the demographics of India i.e. is it truly representative of the Indian population? Why are there different age ranges for men and women? Were data on medications collected? Eg some patients may have treated hypertension with systolic BP <140mmHg….. how are these patients classified? Data is a plural term – please amend text to reflect this. Included variables; would be useful to have more information on this in the text eg how was education/caste/weath status etc categorised? I realise these are broken down in Table 1 but additional information in text would be helpful; in particular on caste as I do not understand the acronyms in Table 1 and no legend is given. Age adjusted rates – could the authors provide more information on how this was conducted please. Was a standard population used and if so what was it. How were urban/rural classifications made? The results section mentions how prevalence is correlated with GDP per state. Where did these data on GDP come from? Results More than half the sample is very young (<30years); is there an explanation for this? Please see my notes above for additional information on the survey. Is the wide disparity in tobacco use between men and women typical? The authors have included two table 1s – please re-label the second as table 2 and amend text accordingly. How did the authors handle men in their standardisation method given that men were included in the survey up to age 54 but standardisation was only done for people aged up to 49yrs (line 50 page 7). Related to this, the figures include variation for the 15-54 years (line 23 page 8) – why a different age bracket for the figures to Table? The authors use the term standardised and adjusted interchangeably; please use standardisation if this is what was done (more information needed in methods). The use of maps is quite nice and presents results quite succinctly. Have the authors considered using multi-level models (or other) to

	try explain inter-state variation? Line 53 page 9 – a different age range given again for graphs? Line 33 page 9 –See my notes above relating to classification of urban/rural and also where did data on GDP come from? Multivariate results: is unusual that smoking was not associated with hypertension. Can the authors provide more information on how the model was made please? Discussion A summary statement is that 10% of the population has hypertension. This seems like an underestimate given global estimates. This underestimate likely results from the young population recruited into survey. Can the authors please address this by comparing to other estimates or using weights perhaps. The authors compare their results to Geldsetzer et al, but do not provide the reader with the key data points form Geldsetzer et al. In other words, what prevalence did Geldsetzer find? There is no discussion of the limitations in this study.

REVIEWER	Bin Liu the Second Hospital of Jilin University, China
REVIEW RETURNED	07-Apr-2019

GENERAL COMMENTS	Ghosh et al analyzed the prevalence of hypertension in India and the related factors. They found that the age-standardized prevalence of hypertension was 11.3%, and was higher in males than in females, in urban adults than in rural adults. The risk factors included obesity, tobacco, and alcohol consumption. The manuscript is basically written fluently, however, there are a few questions: 1. the authors should explain briefly about the random sampling or give a reference in the manuscript even though they didn't conduct the survey. 2. the authors should give the definitions of poorest, poorer, middle, richer, and richest in the wealth quintile catefory. 3. there are two table1, Table 1 sample characteristics, and Table1 prevalence of hypertension in India, 2015-2016. 4. the authors should describe the factors they adjusted for in the multiple logistic regression analysis.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

The paper by titled "Prevalence and associated risk factors of hypertension in India: Evidence from NFHS" was aimed to review the global and interstates prevalences of Hypertension in India" by Ghosh, Soumitra. This a cross section study on data from the 4th round of the National Family Health Survey (NFHS) in India.

The issue is indeed, interesting.

Comment#1

The currents draft is too long requiring be shorten in some specific sections (e.g: Methods could be reduced, Discussion).
Also, number of the tables (4) and figures will likely be recommended sometime legends are scarcely explained or confusing. Also the English language requires revision by an expert.

Response.

Thank you for this comment. Although the method section could not be reduced as the second reviewer suggested us to include additional details of survey data used, we tried to cut down the length of the discussion and results sections. The typos and other language related issues have also been rectified in the revised draft.

Comment#2

The main findings of this study should also (e.g. conclusions) neatly transcribed as: points a)…b)…c)…..d).
In the current draft these key factors are, only maximally mentioned.
To place tables and figure at the end (not in) as a supplement, of the written test.
We would recommended to delete.
Tables: a) Divide Table 1 (the real one in page 7) into 2 parts. (Education as the dividing point). And the now referred as Table 1, really Table 2 (page 8). b) To delete Tables A1 which is too complicated and whose content may essentially described in the text. c) Figures 1,2,4,5 could be suppressed or as complementary supplement.

Response.

Table A1 has been dropped. Besides, as suggested, figures 1, 2, 4 and 5 have been provided as supplementary figures with following names: figure S1, S2, S3 and S4.

Comment#3
We recommended that the selected figures and tables be placed at the end of each page of the manuscript rather than inside the writing text at it appears now.

Response.

As suggested, we have now placed the tables either at the top or end of a page.

We gratefully acknowledge the fact that the strength and limitations /some of them) of this study are recognized explicitly in page 3 (upper part).
Limitations of the study
¬ The use of cross-sectional data that does not allow for exploration of causal pathways underlying the reported associations
¬ The role of behavioural risk factors such as low fruit and vegetable intake and physical inactivity could not be explored in this analysis
¬ Findings are limited to the adults aged between 15 and 54.
Final comment on the contribution of the prevalence of Hypertension to Health in the different population of India. Possible ethnicity and genetic variation between people rural vs urban setting should be better described and better understanding the reported differences found in the prevalence of Hypertension between Northam and Southam states including their respective districts.

Response.

We addressed the above comments, to the extent possible.

Reviewer: 2

Overall, this is an important piece of research that estimates prevalence of hypertension in India at various geographical levels. The authors correctly state that having access to prevalence estimates is

important for managing and measuring the effects of interventions to try help focus efforts on lowering prevalence rates. However, this paper needs substantial revision before it can be published. In brief, it needs much more details on the methods used, in particular how the survey was set up, administered, how the population was recruited, why the age ranges were selected for different genders, why there is a large proportion of young people in the sample. The large proportion of young people is of particular concern as it means the prevalence estimates generated are unlikely to be representative of the general population – indeed at 10% prevalence this is clearly reinforced. Global burden of disease estimates for hypertension are much higher than this. Even with clarity on how the population was set up the authors will need to be very clear on how these estimates generalise to the Indian population. The authors may want to consider weighting to make the estimates more generalisable.
On a technical note there are some typos and some revisions of English may be required.

Comment#1

Abstract
The conclusions are not based on the data presented in the results section i.e. a statement is made about prevalence amongst poor people but there is no explanation for how "poor" is defined in the data or any results presented for "poor" vs "non-poor".

Response. We added the details regarding how households have been defined as poor (low SES) and non-poor (high SES) in the revised version.

Comment#2
Introduction
There are a few acronyms that have not been defined. Please make sure to spell out in first instance of use.
A statement is made referring to inter-state variation in hypertension prevalence, based on differences in deprivation/socio-economic status but no reference is given for either SES or hypertension. Given that the authors also state that this study is the first examination of interstate prevalence I think some clarity here would be helpful; I expect the authors assume differences in hypertension prevalence based on differences in SES etc.

Response. We have now provided expanded form of each acronym the first time it appears in the text.

There are studies that have examined the prevalence of hypertension in some states of India but the results of those studies cannot be generalised for India.

Yes, we presumed that there would be inter-state differences in prevalence of hypertension due to large disparities in social and economic conditions of the populations across states. As suggested by the reviewer, we have now corrected the statement in the revised draft.

Comment#3
Methods
Much more detail is needed in the methods. Regarding the survey, how was the sample derived? How was the survey administered? Were people visited in their homes by trained survey teams? Were teams trained in anthropometric measurements? Was the survey administered via interview or self-complete? How does the respondent rate match the demographics of India i.e. is it truly representative of the Indian population? Why are there different age ranges for men and women? Were data on medications collected? Eg some patients may have treated hypertension with systolic BP <140mmHg….. how are these patients classified?

Response.

National Family Health Survey (NFHS) is the Indian version of Demographic and Health Surveys (DHS) carried out periodically in over 90 countries across the globe. DHS including NFHS are nationally representative household surveys which provide data for a wide range of monitoring and impact evaluation indicators in the areas of population, health and nutrition.

Decisions about the overall sample size required for NFHS-4 were guided by several considerations, paramount among which was the need to produce indicators at the district, state/union territory (UT), and national levels. NFHS 4 was conducted under the stewardship of the Ministry of Health and Family Welfare (MoH&FW), Government of India, and is a collaborative effort of a large number of organisations. ICF, USA, provided technical assistance at all stages of NFHS project. The International Institute for Population Sciences was the nodal agency for the conduction of NFHS 4. Fourteen research organisations, including three Population Research Centres (under MoH&FW, Government of India) shouldered the responsibility of conducting the survey in the different states and UTs of India. The survey used a uniform sample design, questionnaires, field procedures and procedures for biomarker measurements throughout the country to facilitate comparability across states and to ensure the highest possible data quality.

Training of field staff Training was conducted in a tiered fashion. For each of the two fieldwork phases, a Training of Trainers (TOT) course was conducted by IIPS, Mumbai, and ICF. The TOT for the 17 states and union territories included in the first phase (plus the Eastern Region of Uttar Pradesh) was conducted in Puri, Odisha, from 18 August to 2 September 2014, with additional training conducted from 27 October to 3 November 2014 at IIPS, Mumbai. The TOT for the remaining 18 states (plus the Central and Western Regions of Uttar Pradesh) was conducted from 14 November to 1 December 2015 in Chandigarh. The trainees in both TOT courses included project coordinators, health coordinators, statisticians/ demographers, and information technology coordinators from the Field Agencies, and Project Officers/Senior Project Officers from IIPS. The coordinators were responsible for training fieldworkers at the state/UT level. Data collection was conducted in two phases (from 20 January 2015 to 4 December 2016) by 789 field teams. Each team consisted of one field supervisor, three female interviewers, one male interviewer, two health investigators, and a driver. The number of interviewing teams in each state varied according to the sample size.

In each state, interviewers were hired by the selected Field Agencies, taking into consideration their educational background, experience, and other relevant qualifications. Female and male interviewers were assigned to interview respondents of the same sex. In NFHS, the Biomarker Questionnaire covered measurements of height, weight, and haemoglobin for children, and measurements of height, weight, haemoglobin, blood pressure, and random blood glucose for **women age 15-49 and men age 15-54. The different age ranges for men and women** were chosen, keeping in mind the average spousal age gap of 5 years in India.

Data on medications were collected. Sixty-two percent of women and 48 percent of men say that their blood pressure was ever measured prior to the survey: 9 percent of women and 7 percent of men say that on two or more occasions they were told by a doctor or health professional that they have hypertension or high blood pressure. However, only about one-third of diagnosed hypertensives (3% of all women and 2% of all men) are currently taking medicine to lower their blood pressure.

Comment#4

Some patients may have treated hypertension with systolic BP <140mmHg….. how are these patients classified?

Response.

An individual is classified as having hypertension if he/she is currently taking antihypertensive medication to lower his/her blood pressure.

Though a few additional details have been added in the methods section as suggested by the reviewer, not all specifics were incorporated in the revised draft, considering the comments of other two referees with respect to the length of the article.

Comment#5

Data is a plural term – please amend text to reflect this.
Included variables; would be useful to have more information on this in the text eg how was education/caste/wealth status etc categorised? I realise these are broken down in Table 1 but additional information in text would be helpful; in particular on caste as I do not understand the acronyms in Table 1 and no legend is given.
Age adjusted rates – could the authors provide more information on how this was conducted please. Was a standard population used and if so what was it.
How were urban/rural classifications made?
The results section mentions how prevalence is correlated with GDP per state. Where did these data on GDP come from?

Response.

Regarding the mention of data in singular sense at times, it was an inadvertent error, which has been corrected in the revised draft. Additional details regarding the **categorization of education, caste and wealth status** have been provided in accordance with the reviewer's comment (Please see page 6 and 7). We have given more details regarding the calculation of age-adjusted rates (Please see page 6, second para).

The sampling frame of NFHS was based on Census 2011 data. Census has defined geographical areas broadly into two categories-**rural and urban**.

Urban: Constituents of urban are Statutory Towns, Census Towns and Outgrowths.

Statutory Towns: All places with a municipality, corporation, Cantonment Board etc. Whereas, Census town is defined based on the places that satisfy the following criteria: i) a minimum population of 5,000; ii) at least 75 per cent of male working population engaged in non-agricultural pursuits; and a density of population of at least 400 per sq.km. Out Growths are viable units such as a village or part of a village contiguous to a statutory town. Examples of out growths are Railway Colonies, University Campus, port areas, etc.

Rural: All area other than urban are rural. The basic unit for rural areas is the revenue village.

The source of GDP data is Central Statistics Office, Ministry of Statistics and Programme Implementation, Government of India. The same has been mentioned below the figure S4.

Comment #6

Results
More than half the sample is very young (<30years); is there an explanation for this? Please see my notes above for additional information on the survey.
Is the wide disparity in tobacco use between men and women typical?
The authors have included two table 1s – please re-label the second as table 2 and amend text accordingly.
How did the authors handle men in their standardisation method given that men were included in the

survey up to age 54 but standardisation was only done for people aged up to 49yrs (line 50 page 7). Related to this, the figures include variation for the 15-54 years (line 23 page 8) – why a different age bracket for the figures to Table?
The authors use the term standardised and adjusted interchangeably; please use standardisation if this is what was done (more information needed in methods).
The use of maps is quite nice and presents results quite succinctly. Have the authors considered using multi-level models (or other) to try explain inter-state variation?
Line 53 page 9 – a different age range given again for graphs?
Line 33 page 9 –See my notes above relating to classification of urban/rural and also where did data on GDP come from?
Multivariate results: is unusual that smoking was not associated with hypertension. Can the authors provide more information on how the model was made please?

Response.

As per Census 2011 data, **the median age of Indian population** was 24 years and it rose to 26.7 years in 2015 (based on population projections). Therefore, the NFHS data is very much in line with Census figures. As asked by the reviewer, additional details on the survey have been provided earlier.

The wide disparity in tobacco use between men and women is not typical. Traditionally, the prevalence of tobacco consumption among males is much higher than that among females. This may relate to cultural disapproval, prohibiting women from smoking in India. Also, this could be partly due to under-reporting of tobacco-use by women because of social non-acceptance. The prevailing social norms, beliefs, values, and taboo in Indian society are considered to be the major constraints for women to use of tobacco. Below, we have provided the estimates of prevalence of tobacco use by sex from different nationally representative household surveys in India.

	NSSO 50th round	NSSO 52nd round	NFHS-II	NFHS-III	GATS 1	GATS 2
	1993-94	1995-96	1998-99	2005-06	2009-10	2016-17
Gender						
Male	27.9	35.3	29.4	33.4	24.3	19.0
Female	1.9	2.6	2.5	1.4	2.9	2.0

NSSO: National Sample Survey Organisation; NFHS: National Family Health Survey; GATS: Global Adult Tobacco Survey

Table numbers have been changed and the text has also been modified accordingly.

We have responded to the comment on standardization earlier. The analysis is restricted to women and men age 15-49. As suggested by the reviewer, we employed multi-level regression model and the results of the regression analysis have been added in the revised version.

Concerning the comment on the multivariate results for tobacco use, we agree that it is unusual that smoking was not found to be associated with hypertension. In our model, 'tobacco use' is a binary variable with two outcomes. Those who reported use of some form of tobacco including chewing paan (betel leaf), masala or gutkha (smokeless tobacco), smoking cigarette or bidi were assigned '1' and those who reported of not using any form of tobacco were assigned '0'.

However, studies have found similar results earlier. In fact, findings of some studies revealed lower blood pressure levels among smokers compared to ex-smokers (Green, M. S., Jucha, E. & Luz, Y.

1986). Some researchers have also reported lower prevalence of hypertension among the current smokers than among never-smokers and former smokers (Okubo, Y., Miyamoto, T. et al, 2002).

So, the findings are mixed and hence, a few researchers have concluded that it is far from clear the extent to which cigarette smoking is a risk factor for the development of hypertension (Narkiewicz, K., Kjeldsen, S. E. & Hedner, T 2005). However, a recent study examining the life-course impact of smoking on hypertension found no statistically significant association between smoking and the risk of hypertension in the group younger than 35; though smoking was found to be significantly associated with hypertension in the later ages (Gao, Shi and Wang 2017).

One of the possible explanations for not finding 'tobacco use' as a significant predictor of hypertension could be the age composition of our sample. The sample consists of persons aged between 15 and 49. In fact, almost 80% of our sample population are below 40 years.

References

Green, M. S., Jucha, E. & Luz, Y. Blood pressure in smokers and nonsmokers: epidemiologic findings. American heart journal 111, 932–940 (1986).

Okubo, Y., Miyamoto, T., Suwazono, Y., Kobayashi, E. & Nogawa, K. An association between smoking habits and blood pressure in normotensive Japanese men. Journal of human hypertension 16, 91 (2002).

Narkiewicz, K., Kjeldsen, S. E. & Hedner, T. Is smoking a causative factor of hypertension? Blood pressure. 14, 69–71 (2005).

Kaiye Gao, Xin Shi and Wenbin Wang. The life-course impact of smoking on hypertension, myocardial infraction and respiratory diseases. Nature reports (2017)

Comment#7

Discussion

A summary statement is that 10% of the population has hypertension. This seems like an underestimate given global estimates. This underestimate likely results from the young population recruited into survey. Can the authors please address this by comparing to other estimates or using weights perhaps.

The authors compare their results to Geldsetzer et al, but do not provide the reader with the key data points form . In other words, what prevalence did Geldsetzer find?

There is no discussion of the limitations in this study.

Response.

As suggested by the reviewer, I have provided the rate estimated by Geldsetzer et al in the revised version. The estimates of hypertension rate based on NFHS data differ from the prevalence rate calculated by Geldsetzer et al primarily because of two reasons. One, our sample (NFHS) did not include persons aged 50 and above. Two, the survey data used by Geldsetzer et al did not include several major states and UTs of India. Therefore, Geldsetzer et al's estimates can not be generalised for India. These explanations have been provided in detail in the text.

The discussion regarding limitations of the study has been added. Thank you.

Reviewer: 3

Reviewer Name: Bin Liu

Institution and Country: the Second Hospital of Jilin University, China

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

Ghosh et al analyzed the prevalence of hypertension in India and the related factors. They found that the age-standardized prevalence of hypertension was 11.3%, and was higher in males than in females, in urban adults than in rural adults. The risk factors included obesity, tobacco, and alcohol consumption. The manuscript is basically written fluently, however, there are a few questions:
1. the authors should explain briefly about the random sampling or give a reference in the manuscript even though they didn't conduct the survey.
2. the authors should give the definitions of poorest, poorer, middle, richer, and richest in the wealth quintile catefory.
3. there are two table1, Table 1 sample characteristics, and Table1 prevalence of hypertension in India, 2015-2016.
4. the authors should describe the factors they adjusted for in the multiple logistic regression analysis.

Response.

More details regarding the survey and wealth index have been added in the revised manuscript. Please see the methods section, particularly the highlighted portion.

We have taken care of that inadvertent error (regarding table numbers) in the revised version.

As suggested by second reviewer, we have carried out a multilevel logistic regression analysis and the results of the same have been added in the revised write-up.

VERSION 2 – REVIEW

REVIEWER	Sarah-Jo Sinnott LSHTM, UK
REVIEW RETURNED	11-Sep-2019

GENERAL COMMENTS	Thank you for addressing my comments. I have some outstanding clarifications and suggestions. As a constructive top tip, is helpful to give the reviewer line numbers when telling them about the changes you've made. Makes it quick and easy to find the edits. Introduction: the authors have clarified the language on their assumption of varying hypertension prevalence across states, however they still need a reference to support that hypertension depends on SES and backs up their assumption. Methods: the authors have provided information in their response on survey; but this needs to go into manuscript so that the reader can appreciate the methodology used. Please provide a shortened version in manuscript. I cannot see whether the authors have explained the difference in age between men and women in the manuscript – if its not there, please add. Is difficult to find the changes without line numbers in the responses. Please add information on how medication data were collected eg questioned individuals or asked individuals to present their medications. Results/Discussion: the authors explained why the population is so young in their response, thank you for this. This info needs to be added to the manuscript so that the reader can understand how representative the survey data are of the Indian population. Similarly, info on differing tobacco usage being in line with other national estimates should be briefly mentioned in manuscript Discussion: Comparing the results of Geldsetzer and this study: I am not convinced that a 15% difference in prevalence is driven by non-

	inclusion of some states in the Geldsetzer survey. The title of their study tells us they had a nationally representative sample. Can the authors look into the differences a bit better? It may be that the differences are driven solely by age? The authors mentioned rising prevalence of hypertension in India – but not sure they have referenced any studies that show changing trends?

REVIEWER	Bin Liu The Second Hospital of Jilin University, China
REVIEW RETURNED	30-Jul-2019

GENERAL COMMENTS	No further comments.

VERSION 2 – AUTHOR RESPONSE

Reviewer 2

Introduction: the authors have clarified the language on their assumption of varying hypertension prevalence across states, however they still need a reference to support that hypertension depends on SES and backs up their assumption.

Response

I have now the added a reference to qualify the statement. [See page 4, para 3, line 5]

Methods: the authors have provided information in their response on survey; but this needs to go into manuscript so that the reader can appreciate the methodology used. Please provide a shortened version in manuscript. I cannot see whether the authors have explained the difference in age between men and women in the manuscript – if its not there, please add. Is difficult to find the changes without line numbers in the responses. Please add information on how medication data were collected eg questioned individuals or asked individuals to present their medications.

Response

Additional details regarding survey design and data collection process have been incorporated in the revised manuscript. Also, we added the reason for difference in age between men and women in the 'data' section [page 5, last para, last line]. As suggested, information about whether diagnosed hypertensive participants were taking medicine or not has been included [page 6, 1st para, lines 7-8].

Comment

Results/Discussion: the authors explained why the population is so young in their response, thank you for this. This info needs to be added to the manuscript so that the reader can understand how representative the survey data are of the Indian population.

Response

We added the information in the revised manuscript [page 7, last para, lines 2-3].

Comment

Similarly, info on differing tobacco usage being in line with other national estimates should be briefly mentioned in manuscript

Response

We added that too [Page 8, para 1, lines 5-12].

Comment

Discussion: Comparing the results of Geldsetzer and this study: I am not convinced that a 15% difference in prevalence is driven by non-inclusion of some states in the Geldsetzer survey. The title of their study tells us they had a nationally representative sample. Can the authors look into the differences a bit better? It may be that the differences are driven solely by age?

Response

As the reviewer advised, we closely looked at the data used by Geldsetzer and others in their study. The differences could only be explained by the three reasons cited in our paper. However, we agree with the reviewer's observation that the discrepancy could be largely because of age differences between two study samples. So, we have modified the text accordingly in the revised manuscript. [Please see page 14, para 2, line 6, 8-9]

Comment

The authors mentioned rising prevalence of hypertension in India – but not sure they have referenced any studies that show changing trends?

Response

Thank you for the comment. We have added the references of relevant studies. [Page 17, para 3, line 3]