PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Caste, religion and regional differentials in life expectancy at birth in India: cross sectional estimates from recent National Family Health Survey
AUTHORS	Kumari, Meena; Mohanty, Sanjay K

VERSION 1 – REVIEW

REVIEWER	Jair Licio Ferreira Santos Faculdade de Medicina de Ribeirão Preto - USP Brazil
REVIEW RETURNED	25-Nov-2019

GENERAL COMMENTS	Congratulations for the text. Mortality differentials are always a topic that needs to be carefully studied as they can translate important inequalities. Especially in societies where history and cultural tradition may make it difficult to promote equality. I suggest some corrections that, if accepted, could bring better quality to the article. (1) The limitations section of the article should mention the possibility of bias due to the different numbers of people living in the residences. I believe this number varies by caste, income, education, region and other variables. Household surveys can produce selection bias when the number of people per household is varied, due to (a) the greater likelihood of informant absence in small households and (b) differences in instruction by household size. (2) There are two important errors in item 5 in the methods section: premature mortality and mortality in working ages. I have sent the corrections in separate file, due to the difficulty in writing the mathematical expressions in the box comments. (3) Life expectancy refers to a specific age, not an age interval. You say "life expectancy at age 15", not "at age interval 15 - 20". So, tables 2, 3 and 4 must have the first column corrected, showing exact ages, not intervals. Corrections needed for item 5 in Methods. Premature mortality is defined as the chance of dying before the age of 70. That is, the proportion os deaths occurring before 70 : $_0q_{70} = (l_{70} – l_0) / l_0$ And the mortality in working ages is defined as the chance of dying before age 60 among those who reached 15 years of age: $_{15}q_{59} = (l_{15} – l_{59})/l_{15}$

REVIEWER	Marcel Zwahlen Institute of Social and Preventive Medicine University of Bern Switzerland
REVIEW RETURNED	27-Nov-2019

GENERAL COMMENTS	This is an interesting paper reporting on life expectancy (at birth and from other ages) by caste and religion in India. This could only be done by combining information from two data sources: 1) data from National Family Health Survey (NFHS-4), and 2) from the Sample Registration System (SRS). The SRS allows for more robust estimates of age-specific mortality rates, but does not have information on caste and religion. The NFHS-4 is giving individual information on caste and religion but is not designed to estimating precisely and in a representative manner mortality rates. Callibration factors (or in other words : “appropriate” weighting schemes) are used to align the mortality rates from the NFHS-4 data with those from the SRS system. This approach has also been reported for countries from more developed countries, like Switzerland (1). Major comments: 1) The authors state that “The NFHS provides consistent and reliable estimates of fertility, mortality, family planning, child nutritional status, morbidity” This statement about consistent estimates of mortality stands without reference to work that this has been checked. The work in Switzerland has shown, that this statment might not necessarily be true, that individuals included in survey have the same future mortality rates (especially short term) as the whole population (see Moser 2018) even when using the survey weights that are used for making cross-sectional results representative. Authors should provide support for this claim in the introduction or be more precise what they mean by “consistent estimates”. Actually, their additional calibration steps indicate that they needed to do something to align the age-specific rates (table 1). 2) The authors state that “The NFHS was not powered to calculate mortality rate like SRS. To make it comparable with SRS and analyse the mortality data by different socio-economic characteristics we computed multiple calibration factors and link it with the datasets. So, we have not been able to provide the CIs for our estimates.” This is a weak argument: There are certainly ways to obtain 95% confidence intervals. See again the work in (Moser 2018, where they used skew-normal regression, for details see (2)) . But there are other ways of getting uncertainty intervals assuming known calibration factors (weights) via some Monte Carlo or boostsrapping procedure). 3) The explanation how authors arrived at their calibration factors needs a much clearer explanation. The meat of the explanation is in p 4 lines 53 to page 5 lines 6 with a big part of the argument and logic being outsourced to ref 47. Ending with the rather cryptic sentence: “In our calculation C-Factor was calculated by using NFHS and SRS data by sex in India.” Certainly the sentence on p5 , line 18 does not help “Calibration factor is the ratio of response from detector to the analyte

	concentration." This is probably copy paste from a book on calibration of laboratory measurement devices. Minor points: Careful English editing is needed. Just a two examples here P2 , line 57: "The caste structure is also said to have associated with economic wellbeing of the household." Should rather be "be associated" P 5 , line 43: "To calculate annual death rates, the age-specific death, 74945 in total, was divided by 3 to make the death annually, assuming constant mortality rates across the 3 years prior to survey." Rather "the number of age-specific deaths was divided by 3 to estimate an annual death rate, assuming…" – it is hardly "the death" that needs be made annually. 1. Moser A, et al. Calibration adjustments to address bias in mortality analyses due to informative sampling-a census-linked survey analysis in Switzerland. PeerJ 2018;6:e4376. 2. Moser A, e tal. Modeling absolute differences in life expectancy with a censored skew-normal regression approach. PeerJ 2015;3:e1162.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Jair Licio Ferreira Santos
Institution and Country: Faculdade de Medicina de Ribeirão Preto - USP, Brazil
Please state any competing interests or state 'None declared': None
Reply:
Thank you for your suggestion. We have modified it Competing Interests: 'None declared'.

Congratulations for the text. Mortality differentials are always a topic that needs to be carefully studied as they can translate important inequalities. Especially in societies where history and cultural tradition may make it difficult to promote equality. I suggest some corrections that, if accepted, could bring better quality to the article.

(1) The limitations section of the article should mention the possibility of bias due to the different numbers of people living in the residences. I believe this number varies by caste, income, education, region and other variables. Household surveys can produce selection bias when the number of people per household is varied, due to (a) the greater likelihood of informant absence in small households and (b) differences in instruction by household size.
Reply:

We thank the reviewer for this suggestion. We understand that reviewer intends to mention the missing information on deaths in surveys. We have added this point in limitation. Besides, the survey did not collect data on cause of death and sample size was not adequate for providing estimates for all states of India. All these points are added in limitation section.
The adjustment for missing death cases was done as follows:
The number of missing cases was 479 that accounts 0.6 percent of total deaths. Hence the missing data may not affect the estimates largely. However, we have adjusted the missing cases by distributing 479 cases equally in each age group.

(2) There are two important errors in item 5 in the methods section: premature mortality and mortality in working ages. I have sent the corrections in a separate file, due to the difficulty in writing the mathematical expressions in the box comments.

Reply: We are thankful to the reviewer for this suggestion. However, the premature mortality is defined as the deaths occur before the age 70 years, according world health organization.
This is the common description of life table from Demography: measuring and modeling population processes [1].
The abridged life table is constructed to estimate the life expectancy at birth, adult mortality (15q_59) and premature mortality (0q_70) by caste, religion and regions. The premature mortality is defined as any death under 70 years of age. The mathematical form of Premature (0q_70) = (l_0-l_70)/l_0 ; and Mortality in working age (15q_59) = (l_15-l_59)/l_15 .

(3) Life expectancy refers to a specific age, not an age interval. You say "life expectancy at age 15", not "at age interval 15 - 20". So, tables 2, 3 and 4 must have the first column corrected, showing exact ages, not intervals.

Reply: Thank you so much. We have changed the first column of table 2, 3 and 4 to exact age.

Reviewer: 2
Reviewer Name: Marcel Zwahlen
Institution and Country: Institute of Social and Preventive Medicine, University of Bern, Switzerland
Please state any competing interests or state ‘None declared’

Reply: We have stated the competing interest as “None declared”.

This is an interesting paper reporting on life expectancy (at birth and from other ages) by caste and religion in India. This could only be done by combining information from two data sources: 1) data from National Family Health Survey (NFHS-4), and 2) from the Sample Registration System (SRS). The SRS allows for more robust estimates of age-specific mortality rates, but does not have information on caste and religion. The NFHS-4 is giving individual information on caste and religion but is not designed to estimating precisely and in a representative manner mortality rates.
Calibration factors (or in other words: “appropriate” weighting schemes) are used to align the mortality rates from the NFHS-4 data with those from the SRS system. This approach has also been reported for countries from more developed countries, like Switzerland (1).
Major comments:

1)
The authors state that “The NFHS provides consistent and reliable estimates of fertility, mortality, family planning, child nutritional status, morbidity”
This statement about consistent estimates of mortality stands without reference to work that this has been checked. The work in Switzerland has shown, that this statement might not necessarily be true, that individuals included in survey have the same future mortality rates (especially short term) as the whole population (see Moser 2018) even when using the survey weights that are used for making cross-sectional results representative.
Authors should provide support for this claim in the introduction or be more precise what they mean by “consistent estimates”. Actually, their additional calibration steps indicate that they needed to do something to align the age-specific rates (table 1).

Reply: Thank you for this suggestion. We actually intended to mention the reliability of estimates and not that of consistency. We have revised the sentences as:
“The NFHS provides reliable estimates of fertility, child mortality, family planning, child nutritional status and childhood morbidity for the country. With regard to mortality, NFHS 1, 2 and 3 were intended to provide reliable estimates of infant and child mortality only. However, in NFHS 4, an attempt was made to capture the overall mortality.”

2)
The authors state that “The NFHS was not powered to calculate mortality rate like SRS. To make it comparable with SRS and analyse the mortality data by different socio-economic characteristics we computed multiple calibration factors and link it with the datasets. So, we have not been able to provide the CIs for our estimates.”
This is a weak argument: There are certainly ways to obtain 95% confidence intervals. See again the work in (Moser 2018, where they used skew-normal regression, for details see (2)) . But there are other ways of getting uncertainty intervals assuming known calibration factors (weights) via some Monte Carlo or bootsrapping procedure).

Reply: We thank the reviewer for this suggestion. The estimate of life expectancy was derived using life table methods. In life table method there no provision of calculating confidence interval.
3)
The explanation how authors arrived at their calibration factors needs a much clearer explanation. The meat of the explanation is in p 4 lines 53 to page 5 lines 6 with a big part of the argument and logic being outsourced to ref 47. Ending with the rather cryptic sentence: “In our calculation C-Factor was calculated by using NFHS and SRS data by sex in India.”
Certainly the sentence on p5 , line 18 does not help “Calibration factor is the ratio of response from detector to the analyte concentration.” This is probably copy paste from a book on calibration of laboratory measurement devices.

Reply: In the revised draft we have renamed “Calibration factor” as “Adjustment factor” since the calibration factor as a terminology is usually used in life sciences and might introduce some misconceptions while using in case of social sciences. Also, we have provided a proper explanation for calculation of the adjustment factor along with mathematical expression in the draft file. Further, we have also revised the sentence.

Mathematically:
A0-1 (Male, Female, Total) = (M_(0-1 (Male,Female,Total))^SRS)/(M_(0-1 (Male,Female,Total))^NFHS) ---------- (1)
A1-4 (Male, Female, Total) = (M_(1-4 (Male,Female,Total))^SRS)/(M_(1-4 (Male,Female,Total))^NFHS) ---------- (2)
.
.
.
A85+ (Male, Female, Total) = (M_(85+(Male,Female,Total))^SRS)/(M_(85+ (Male,Female,Total))^NFHS) ---------- (3)

Similarly to calculate the adjustment factor for each age group 0-5, 5-10…..80-85 years, now by using this adjustment factor to calculate mortality rate for each age group.

M_(0-1(Male,Female,Total))^SRS= A_(0-1(Male,Female,Total))*M_(0-1(Male,Female,Total))^NFHS---------- (4)

M_(1-4 (Male,Female,Total))^SRS= A_(1-4 (Male,Female,Total))*M_(1-5 (Male,Female,Total)
)^NFHS---------- (5)
.
.
.

M_(85+ (Male,Female,Total))^SRS= A_(85+ (Male,Female,Total))*M_(85+ (Male,Female,Total)
)^NFHS---------- (6)

In this calculation 'M' signifies the mortality rate for the given subgroup from the national Family Health Survey (NFHS) and Simple Registration system (SRS) dataset and 'A' signifies the subgroup specific adjustment factor for each age group. We then assumed this adjustment factor to be constant across the each caste, religion, and region differentials within this subgroup.

Minor points:
Careful English editing is needed.

Just a two examples here
P2 , line 57: "The caste structure is also said to have associated with economic wellbeing of the household." Should rather be "be associated"
P 5 , line 43: "To calculate annual death rates, the age-specific death, 74945 in total, was divided by 3 to make the death annually, assuming constant mortality rates across the 3 years prior to survey."

Rather "the number of age-specific deaths was divided by 3 to estimate an annual death rate, assuming…" – it is hardly "the death" that needs be made annually.

1. Moser A, et al. Calibration adjustments to address bias in mortality analyses due to informative sampling-a census-linked survey analysis in Switzerland. PeerJ 2018;6:e4376.
2. Moser A, e tal. Modeling absolute differences in life expectancy with a censored skew-normal regression approach. PeerJ 2015;3:e1162

1 Preston Samuel H, Patrick H, Michel G. Demography: measuring and modeling population processes. MA Blackwell Publ 2001.

VERSION 2 – REVIEW

REVIEWER	Jair Licio Ferreira Santos Faculdade de Medicina de Ribeirão Preto – USP
REVIEW RETURNED	13-Feb-2020

GENERAL COMMENTS	Congratulations for your effort. The paper has achieved greater quality after your corrections,

REVIEWER	Marcel Zwahlen Institute of Social and Preventive Medicine, University of Bern, Switzerland
REVIEW RETURNED	16-Feb-2020

GENERAL COMMENTS	The authors accounted for many of the comments and the revised manuscript is now much clearer. Thank you.

	However, I am very much surprised by definite response they give on the possibility to provide 95% confidence intervals for the reported life expectancies. I quote from their response “In life table method there no provision of calculating confidence interval.” They clearly do not know the relevant literature. One of the first method to do this was published 60 years ago !! Chiang, C.L. (1960). A stochastic study of the life table and its applications: II. Sample variance of the observed expectation of life and other biometric functions. Human Biology 32(3): 221–238. And of course it was also elaborated in more details in the book “Chiang, C.L. (1984). The life table and its applications. Malabar: Robert E. Krieger.” There are also freely available spreadsheet implementations: https://www.demogr.mpg.de/en/publications_databases_6118/publications_1904/mpidr_technical_reports/spreadsheet_for_calculation_of_confidence_limits_for_any_life_table_or_healthy_life_table_quantity_3853/ Clearly, I suggest that the authors go ahead and provide appropriately calculated uncertainty intervals.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Jair Licio Ferreira Santos
Institution and Country: Faculdade de Medicina de Ribeirão Preto - USP
Please state any competing interests or state ‘None declared’: none

Reply:
Thank you for your suggestion. We have modified it Competing Interests: ‘None declared’: None

Reviewer: 2
Reviewer Name: Marcel Zwahlen
Institution and Country:
Institute of Social and Preventive Medicine, University of
Bern, Switzerland
Please state any competing interests or state ‘None declared’: none

Please leave your comments for the authors below
Reply:
Thank you for your suggestion. We have modified it Competing Interests: ‘None declared’: None

The authors accounted for many of the comments and the revised manuscript is now much clearer. Thank you.
However, I am very much surprised by definite response they give on the possibility to provide 95% confidence intervals for the reported life expectancies. I quote from their response “In life table method there no provision of calculating confidence interval.”
They clearly do not know the relevant literature. One of the first method to do this was published 60 years ago !!
Chiang, C.L. (1960). A stochastic study of the life table and its applications: II. Sample variance of the

observed expectation of life and other biometric functions. Human Biology 32(3): 221–238. And of course it was also elaborated in more details in the book “Chiang, C.L. (1984). The life table and its applications. Malabar: Robert E. Krieger.”
There are also freely available spreadsheet implementations:
https://www.demogr.mpg.de/en/publications_databases_6118/publications_1904/mpidr_technical_reports/spreadsheet_for_calculation_of_confidence_limits_for_any_life_table_or_healthy_life_table_quantity_3853/

Clearly, I suggest that the authors go ahead and provide appropriately calculated uncertainty intervals.
Reply:
We thank the reviewer for this suggestion. The suggestion were valid and were attended. The CIs are shown in appendices.

VERSION 3 – REVIEW

REVIEWER	Marcel Zwahlen Institute of Social and Preventive Medicine, University of Bern, Switzerland
REVIEW RETURNED	27-Apr-2020

GENERAL COMMENTS	In this second revision the authors accounted again for many of the comments and now provide 95% CI for the life expectancy estimates. Thank you. However, the authors should double check how they calculated or reported their 95%CIs, as in the abstract they provide estimates that are outside of the 95% CI as in the key sentence of LE according to religion: “Life expectancy at birth was higher among females than among males across social groups in India. It was higher among Christians 68.1 years [95% CI: 69.81-74.30], followed by Muslims 66.0 years [95% CI: 66.71-68.45] and Hindus 65.0 years [95% CI: 67.34-68.09]” In the results section in the manuscript, other numbers are reported. Furthermore, they should give in the methods section a reference or weblink indicating which method they used to calculate the 95% CI. They should easily be able to make these small revisions.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Marcel Zwahlen
Institution and Country: Institute of Social and Preventive Medicine, University of Bern, Switzerland

In this second revision the authors accounted again for many of the comments and now provide 95% CI for the life expectancy estimates.
Reply: Thank you.
Please state any competing interests or state ‘None declared’: none
Reply: Thank you for your suggestion. We have modified it Competing Interests: ‘None declared’: None

However, the authors should double check how they calculated or reported their 95%CIs, as in the abstract they provide estimates that are outside of the 95% CI as in the key sentence of LE according to religion:
“Life expectancy at birth was higher among females than among males across social groups in India. It was higher among Christians 68.1 years [95% CI: 69.81-74.30], followed by Muslims 66.0 years [95% CI: 66.71-68.45] and Hindus 65.0 years [95% CI: 67.34-68.09]”
In the results section in the manuscript, other numbers are reported.
Reply: I really thank you for insightful suggestion and bringing the typo error into notice. The error was corrected in revised manuscript.

Suggestion: Furthermore, they should give in the methods section a reference or weblink indicating which method they used to calculate the 95% CI.
They should easily be able to make these small revisions.

Reply: We are sorry for that, the suggestion are included. We have added the reference
Andreev EM, Shkolnikov VM. Spreadsheet for calculation of confidence limits for any life table or healthy-life table quantity. Rostock: Max Planck Inst Demographic Research (MPIDR) Technical Report- 2010;5.

VERSION 4 – REVIEW

REVIEWER	Marcel Zwahlen Institute of Social and Preventive Medicine, University of Bern, Switzerland
REVIEW RETURNED	29-May-2020

GENERAL COMMENTS	Remaining issues after revision 2 have been addressed to my satisfaction. No further comments from my side.