PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Urban-rural lifespan disparities and cause-deleted analysis: Evidence from China
AUTHORS	Chen, Mengxue; Canudas-Romo, Vladimir

VERSION 1 – REVIEW

REVIEWER	van Raalte, Alyson Max Planck Institute for Demographic Research, Survival and Longevity
REVIEW RETURNED	14-May-2021

GENERAL COMMENTS	This paper is one of the most comprehensive papers that I have seen examining age patterns of mortality in China. The authors look at these differences by sex and rural/urban residence, and how they have changed over time using state of the art demographic techniques. These are important questions. Too much of the literature is based on European and North American populations and it remains unclear the extent to which we should expect other populations to follow the same epidemiologic patterns. As such, this paper is timely and important. Overall, I like the paper, but my main concern is that I would like to see the authors give more consideration to the underlying data quality and how this might impact their results. I'm certainly not an expert on Chinese age patterns of mortality or Chinese data sources, but looking at the mortality curves in the appendix gives me pause. Older ages: The authors acknowledge data problems at older ages, which they mitigate by truncating the data at ages 85+. I worry that this decision makes it more challenge to interpret differences in lifespan variation between the population groups, because of differences in survival to age 85. There's a large literature on how truncation/censoring impacts levels and trends of lifespan variation (e.g. (Engelman, Canudas-Romo, & Agree, 2010; Robine, 2001)). While these mostly looked at the left tail, I would assume that the right tail might also be important, since variability is overall fairly sensitive to differences in mortality at the tails, particularly compared to life expectancy. Given the differences in survival to age 85, why not consider extrapolating mortality for these ages, for instance by a parametric model like the Kannisto model? True, cause of death analysis would require an assumption that either cause of death proportions over these ages are the same as those observed at ages 80-84, or perhaps, given the problems with multimorbidity and assessing underlying causes at older ages anyway, the cause-specific impact could be looked at only up to ages 75 or so.

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

	Younger ages: The lower mortality over ages 20-24 compared to ages 15-19 and 10-14 among females is suspicious. I had a look at all lifetables from Japan, Taiwan, Hong Kong, and Sweden contained in the Human Mortality Database to see how common this was. First, comparing 20-24 with 15-19: on average over all 5x1 lifetables, the mortality rate ratio (Mx[20-24]/Mx[15-19]) was 1.24 (TWN), 1.58 (JPN), 1.39 (HKG), and 1.21 (SWE). The number of times where mortality at 20-24 was lower than 15-19 was: 0/50 (TWN), 0/73 (JPN), 2/32 (HKG), and 8/269 (SWE). Second, on average over all 5x1 lifetables, the mortality rate ratio (Mx[20-24]/Mx[10-14]) was 2.33 (TWN), 2.90 (JPN), 2.23 (HKG), and 1.47 (SWE). The number of times where mortality at 20-24 was lower than 10-14 was: 0/50 (TWN), 0/73 (JPN), 0/32 (HKG), and 11/269 (SWE, all before 1858 when infectious disease mortality was high, and prior to antibiotics and vaccines). Why China should have such a radically different pattern of mortality when children come of age from the rest of Asia and the gold-standard Swedish data should either be explained, or if there is no easy explanation, treated with suspicion. For more background, in 2019, I attended the HMD Mortality symposium where there were several talks on China's unusual age pattern of reported mortality compared to all other countries known to have good data (some of the talks are linked below). Others also found unusual Chinese patterns in the youngest age groups, where concerns were openly aired that the penalties in existence for exceeding one or two children meant that some individuals were not registered at birth, but were showing up in the surveillance systems at older childhood or early adult ages, distorting the age pattern of mortality up to early adulthood. Some of the published literature on this topic can be found in the slides of the following talks from this workshop: https://www.mortality.org/Public/HMD_5th_Symposium/Public/Gu_pres.pdf https://www.mortality.org/Public/HMD_5th_Symposium/Public/Li_Mi_pres.pdf https://www.mortality.org/Public/HMD_5th_Symposium/Public/Cai_pres.pdf https://www.mortality.org/Public/HMD_5th_Symposium/Private/Missov_Nemeth_Li_et.al..pdf Overall, my data quality concerns are actually stronger for these younger ages than for the older ages, especially since lifespan disparity is most sensitive to mortality differences at younger ages. I would urge the authors to consider adjusting the all-cause mortality baseline using suitable existing indirect demographic methods or by making use of the known empirical regularities to the age patterns of mortality across populations with similar mortality levels through adulthood, where the underlying data quality is less in question. Given these known data quality issues, it might also make sense to give a range of plausible model-based estimates of life expectancy and lifespan disparity rather than a single point estimate. Alternatively, consider truncation to adulthood. Other comments: 1. The discussion could be strengthened by bringing in a broader literature on age patterns of mortality and the unfolding epidemiologic transition in different parts of the world, to give the reader a better sense of how to interpret these findings and the

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

	extent to which they are unexpected. Right now the discussion is rather descriptive and the case for the importance of these results isn't as strong as it could be. 2. Paragraph, Line 259: It would be good to draw on the work of (Seligman, Greenberg, & Tuljapurkar, 2016), who examined this specifically. 3. It was an interesting choice to look at cause-deleted life tables rather than a full age and cause of death decomposition, particularly when looking at the age components in the appendix. Do they differ much? What was the justification for using one strategy over another? Was this to make it easier to see the impact of each cause simultaneously on life expectancy and lifespan disparity? 4. Line 76-77: I love this article, but it is more about a new life expectancy decomposition technique than about measuring dispersion of length of life in a population (at least that's not the main purpose of the article). 5. Comments on the figures: (a) Why the reverse scale for lifespan disparity? I've never seen that and it threw me off at first, when I imagined that China was experiencing radically different patterns. (b) The legend is a little busy. Since there is a lot of white space it might be easier to make better use of direct labeling on the figure so that the reader doesn't have to look back and forth from the figure to the legend. References Engelman, M., Canudas-Romo, V., & Agree, E. M. (2010). The Implications of Increased Survivorship for Mortality Variation in Aging Populations. Population and Development Review, 36(3), 511-539. Robine, J.-M. (2001). Redefining the Stages of the Epidemiological Transition by a Study of the Dispersion of Life Spans: The Case of France. Population: An English Selection, 13(1), 173-193. Seligman, B., Greenberg, G., & Tuljapurkar, S. (2016). Equity and length of lifespan are not the same. Proceedings of the National Academy of Sciences, 113(30), 8420-8423. doi:10.1073/pnas.1601112113

REVIEWER	Sauerberg, Markus Austrian Academy of Sciences Vienna Institute of Demography For full transparency, I would like to mention that I am involved in a research project with the second author of the study
REVIEW RETURNED	03-Jun-2021

GENERAL COMMENTS	The study „Urban-rural lifespan disparities and cause-specific contributions: Evidence from China" includes an extensive empirical analysis, addressing urban-rural as well as male-female gaps in life expectancy and lifespan disparities, and quantifies the contribution made by causes of death. Accordingly, the research refers to several important topics, making it relevant for a broad readership. I enjoyed reading the clearly-written manuscript and appreciated all the useful information (on method, data, and figures) provided by the authors in the appendix. From my point of view, the study is almost ready for publication. My suggestion for further improvement is adding more information on the interpretation of results in the discussion section. Health and mortality (especially when data on specific causes of death is available) can be attributed to past health behaviours,

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

	risk, and exposures of individuals. These will likely differ between women and men as well as between urban and rural regions. On page 14 (first paragraph), the authors discuss their findings on lung and liver cancer in terms of gender differences in smoking and drinking (which is great!). I would welcome more discussion in this direction. For example, what about external causes? On page 11 (third paragraph), the authors describe that external causes have a stronger influence on life expectancy and lifespan disparity for men. Could this be connected to differences in occupational hazards (risks of accidents) between women and men (see e.g., Zhu et al. 2019 or He et al. 2005)? How do the findings relate to differences in urban and rural regions? I could imagine that pollution in urban areas might play a role (e.g., Somboonsin and Canudas-Romo 2021)? I believe that adding some information in this regard would be extremely useful to the readers. I also have some minor comments and suggestions that I hope can improve the paper: - The definition of urban and rural (page 6, data source section): I was wondering about the classification, i.e., based on what do we distinguish between urban and rural China? - The authors calculated life expectancy and lifespan disparity with life tables for ages between 0 and 85. I agree with the approach (given the mentioned data limitations), but I was wondering about the terminology. Wouldn't it be more appropriate to refer to truncated (or partial) life expectancy and truncated (or partial) lifespan disparity as you do on page 7 "Measure of lifespan inequality")? Please note that I would not recommend rerunning all the empirical analyses just to get the ggplots with different labels. The current labels with "(Age 0-85)" making it already very clear to the reader to which age span the measures refer. - Speaking about terminology, I could not find the explanation for CVD in the manuscript. I guess this means cardiovascular disease? I would suggest mentioning it once in the text and using the abbreviation throughout the paper (not switching between the abbreviation and written-out version). - Limitations of mortality data: The authors mention the possibility of underreporting. This reminded me of the discussion on quality of cause of death statistics in rural China (e.g., Wang et al. 2007), which also relates to the topic of this study. This might be another potential data restriction worthwhile for consideration. - The authors note the issue of competing risks of death for cause-deleted life table analyses. The assumption of independent causes of death is likely to be more problematic for some diseases than for others, e.g., external causes can be considered as more independent than CVD? Maybe this can be mentioned so readers are aware which conclusions should be drawn with more caution and which results are less affected by the assumption. - The equation 4 on page 28 is equal to equation 2 on the previous page, no? - In Appendix 4, the authors provide a country-comparison, including estimates for Australia, Norway, and the US. I could imagine that these great efforts would remain unseen and underappreciated if not mentioned more directly in the main text. - Appendix 6: The provided graphs are very helpful. I was only wondering whether the figure caption "death distribution" should be replaced with age-specific mortality rates? The plots show death rates on a logarithmic scale ("log_mx"), no? - Appendix 7: This is an interesting description (by the way, "the impact a certain cause…" the "of" seems to be missing). After

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

	reading it and looking at the e-dagger equation, I had the following question: Potential gains in life expectancy are not only affected by the number of deaths from a specific cause but also by the age at death. Averting deaths at younger ages will save a larger number of person-years lived and result in relatively large increases in remaining life expectancy. Further, lifespan disparity is affected by both, remaining life expectancy at age x and the age distribution of deaths. Will lifespan disparity necessarily decrease more when causes of deaths that mostly occur at young ages are deleted compared to deleting causes of death that occur more at older ages? Or can the effect be compensated because averting deaths at young ages leads to gains in longevity, making e-dagger larger for the remaining causes of deaths? My question relates also to the discussion on the age distribution of deaths and the threshold age (page 14). Does it mean that deleting causes of death occurring below this threshold age will always translate into life expectancy increases as well as in lifespan disparity decreases? As I stated above, this is only minor, but I got curious about the author's opinion on the relationship between gains in longevity and changes in lifespan disparity in terms of deleting specific causes of deaths at different ages. References He J, Gu D, Wu X, Reynolds K, Duan X, Yao C, Wang J, Chen CS, Chen J, Wildman RP, Klag MJ, Whelton PK. Major causes of death among men and women in China. N Engl J Med. 2005 Sep 15;353(11):1124-34. Somboonsin P, Canudas-Romo V. Mortality attributable to fine particulate matter in Asia, 2000–2015: a cross-sectional cause-of-death analysis. BMJ Open 2021;11. Wang L, Yang G, Jiemin M, et al. Evaluation of the quality of cause of death statistics in rural China using verbal autopsies. J Epidemiol Community Health. 2007;61(6):519-526. Zhu, J., Cui, L., Wang, K. et al. Mortality pattern trends and disparities among Chinese from 2004 to 2016. BMC Public Health 19, 780 (2019).

REVIEWER	Nigri, Andrea University of Foggia
REVIEW RETURNED	26-Jun-2021

GENERAL COMMENTS	This interesting manuscript analyses the urban-rural gaps in life expectancy and lifespan disparity by sex in China (2006–2016), quantifying the contributions made by causes of death. Using data from the Chinese Disease Surveillance Point System, the authors' findings support how rural residents are quickly catching up to their urban counterparts, shrinking the gap in both life expectancy and lifespan disparity. On the contrary, the gender gap remains large. The manuscript is well-written and my comments are few. To their credit, the authors were aware of possible main limitations (e.g. underreporting and unadjusted competing risks of death), acknowledging them in the manuscript. 1. In my opinion, the abstract does not express this paper's potential. Therefore, I suggest briefly expand the abstract's Objectives section to better introduce the hypothesis (research

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

	questions or prior beliefs), and thus what the authors address through this investigation. This also applies to the Introduction section, which should be enriched with this same content. Similarly, as a service to readers, please briefly expand the Conclusions section (in the abstract) by introducing some notes about the contribution of causes of death. 2. Though life expectancy is well-known, lifespan may be less so. As such, I recommend motivating the choice of an e-dagger to measure the dispersion in life. 3. Among the Strengths and limitations, the authors state that adjustment measures were taken to minimize the impact of underreporting. I think this may sound like an overstatement since the authors chose to simply remove the data of people aged over 85. 4. Following my previous point, China's data suffers from important issues of birth underreporting and an incorrect estimate of infant mortality. Though the literature provides some insights on this topic (Li et al., 2018; Merli & Raftery, 2000), in my understanding, the authors do not address this issue in the paper. 5. The Discussion section provides the authors' findings, where they state, "After analysing the change of life expectancy and lifespan disparity of the subpopulations in China from 2006 to 2016, we find that the increase of life expectancy in China is accompanied by a reduction in the level of dispersion in lifespan, which is consistent with what has been found in countries...." Looking at Figure 1, the reader may clearly find an absence of correlation for the urban male population. This proves to be more problematic in Figure 2 with the case of CVD in urban females and the minor impact of respiratory diseases in urban males. Might this aspect be a result of the authors' choice not to adjust competing risks of death? Can the Authors briefly argue and defend their choice? Furthermore, I ask the authors to do their best to improve (if possible) Figure 2, making it easier to read. Literature - The estimation of death underreporting in the 2010 population census based on DCMD model life tables. Li Cheng, Mi Hong, Sun Lingxue. Population Research Vol. 42, No. 02, 99–112, March 2018. - Are births underreported in rural China? Manipulation of statistical records in response to China's population policies. M. Giovanna Merli; Adrian E. Raftery. Demography (2000) 37 (1): 109–126.

VERSION 1 – AUTHOR RESPONSE

> Reviewer: 1

> Ms. Alyson van Raalte, Max Planck Institute for Demographic Research, Erasmus University Medical Centre

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

> Comments to the Author:
> Review: Urban-rural lifespan disparities and cause-specific contributions: Evidence from China
>
> This paper is one of the most comprehensive papers that I have seen examining age patterns of mortality in China. The authors look at these differences by sex and rural/urban residence, and how they have changed over time using state of the art demographic techniques. These are important questions. Too much of the literature is based on European and North American populations and it remains unclear the extent to which we should expect other populations to follow the same epidemiologic patterns. As such, this paper is timely and important.
Author response: Thank you very much for your encouraging words and helpful advice. We've revised the paper based on your comments and addressed the issues regarding the data quality by taking more adjustment measures.
> Overall, I like the paper, but my main concern is that I would like to see the authors give more consideration to the underlying data quality and how this might impact their results. I'm certainly not an expert on Chinese age patterns of mortality or Chinese data sources, but looking at the mortality curves in the appendix gives me pause.
> Older ages: The authors acknowledge data problems at older ages, which they mitigate by truncating the data at ages 85+. I worry that this decision makes it more challenge to interpret differences in lifespan variation between the population groups, because of differences in survival to age 85. There's a large literature on how truncation/censoring impacts levels and trends of lifespan variation (e.g. (Engelman, Canudas-Romo, & Agree, 2010; Robine, 2001)). While these mostly looked at the left tail, I would assume that the right tail might also be important, since variability is overall fairly sensitive to differences in mortality at the tails, particularly compared to life expectancy.
> Given the differences in survival to age 85, why not consider extrapolating mortality for these ages, for instance by a parametric model like the Kannisto model? True, cause of death analysis would require an assumption that either cause of death proportions over these ages are the same as those observed at ages 80-84, or perhaps, given the problems with multimorbidity and assessing underlying causes at older ages anyway, the cause-specific impact could be looked at only up to ages 75 or so.
Author response: Thank you for your advice, we've followed your suggestions and now use the Kannisto model to extrapolate the old-age mortality to age 110+ instead of removing the data of people aged above 85. Please see line 108-114 on page 7 for our new adjustment measures, and also Appendix 1 for data before and after adjustment.
> Younger ages: The lower mortality over ages 20-24 compared to ages 15-19 and 10-14 among females is suspicious. I had a look at all lifetables from Japan, Taiwan, Hong Kong, and Sweden contained in the Human Mortality Database to see how common this was.
> First, comparing 20-24 with 15-19: on average over all 5x1 lifetables, the mortality rate ratio (Mx[20-24]/Mx[15-19]) was 1.24 (TWN), 1.58 (JPN), 1.39 (HKG), and 1.21 (SWE). The number of times where mortality at 20-24 was lower than 15-19 was: 0/50 (TWN), 0/73 (JPN), 2/32 (HKG), and 8/269 (SWE).
> Second, on average over all 5x1 lifetables, the mortality rate ratio (Mx[20-24]/Mx[10-14]) was 2.33 (TWN), 2.90 (JPN), 2.23 (HKG), and 1.47 (SWE). The number of times where mortality at 20-24 was lower than 10-14 was: 0/50 (TWN), 0/73 (JPN), 0/32 (HKG), and 11/269 (SWE, all before 1858 when infectious disease mortality was high, and prior to antibiotics and vaccines).
> Why China should have such a radically different pattern of mortality when children come of age from the rest of Asia and the gold-standard Swedish data should either be explained, or if there is no easy explanation, treated with suspicion.
> For more background, in 2019, I attended the HMD Mortality symposium where there were several talks on China's unusual age pattern of reported mortality compared to all other countries known to have good data (some of the talks are linked below). Others also found unusual Chinese patterns in the youngest age groups, where concerns were openly aired that the penalties in existence for exceeding one or two children meant that some individuals were not registered at

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

birth, but were showing up in the surveillance systems at older childhood or early adult ages, distorting the age pattern of mortality up to early adulthood. Some of the published literature on this topic can be found in the slides of the following talks from this workshop:
>
https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.mortality.org%2FPublic%2FHMD_5th_Symposium%2FPublic%2FGu_pres.pdf&data=04%7C01%7CMengxue.Chen%40a
nu.edu.au%7C1f7a1b4824924404563808d978a5c4c2%7Ce37d725cab5c46249ae5f0533e486437%7C
0%7C0%7C637673475624332416%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoi
V2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=%2FU3XRWMIJ8jVLRZQ3QcgS%2
FgRFBzS2FzFWMgNaGEm0h4%3D&reserved=0
>
https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.mortality.org%2FPublic%2FHMD_5th_Symposium%2FPublic%2FLi_Mi_pres.pdf&data=04%7C01%7CMengxue.Chen%4
0anu.edu.au%7C1f7a1b4824924404563808d978a5c4c2%7Ce37d725cab5c46249ae5f0533e486437%
7C0%7C0%7C637673475624332416%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQI
joiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=K64vsCM%2F8ZANmn%2Bt81z
Vw2NTyTcq7bNBNCNh7N9NDcU%3D&reserved=0
>
https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.mortality.org%2FPublic%2FHMD_5th_Symposium%2FPublic%2FCai_pres.pdf&data=04%7C01%7CMengxue.Chen%40a
nu.edu.au%7C1f7a1b4824924404563808d978a5c4c2%7Ce37d725cab5c46249ae5f0533e486437%7C
0%7C0%7C637673475624332416%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoi
V2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=MEzvFVfIA%2FLj1nWEAZhlGRgp
vsys2dH5I7nqMd1SWUM%3D&reserved=0
>
https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.mortality.org%2FPublic%2FHMD_5th_Symposium%2FPrivate%2FMissov_Nemeth_Li_et.al..pdf&data=04%7C01%7CMe
ngxue.Chen%40anu.edu.au%7C1f7a1b4824924404563808d978a5c4c2%7Ce37d725cab5c46249ae5f
0533e486437%7C0%7C0%7C637673475624332416%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4w
LjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=kiaMh%2BSUd
RDnRzN5OLyk1gDwpcCbfhu4DJcVbYtEFEo%3D&reserved=0
> Overall, my data quality concerns are actually stronger for these younger ages than for the older ages, especially since lifespan disparity is most sensitive to mortality differences at younger ages. I would urge the authors to consider adjusting the all-cause mortality baseline using suitable existing indirect demographic methods or by making use of the known empirical regularities to the age patterns of mortality across populations with similar mortality levels through adulthood, where the

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

underlying data quality is less in question. Given these known data quality issues, it might also make sense to give a range of plausible model-based estimates of life expectancy and lifespan disparity rather than a single point estimate. Alternatively, consider truncation to adulthood.
>
Author response:
Thank you so much for taking the time to check the HMD database, make the calculation, and share with us the slides from the HMD Mortality symposium. The information in your comments and in the slides are really helpful. We appreciate your concerns about the data quality concerns (especially among the younger age groups), and made some revisions to address these issues.
Following your advice, we
1) addressed the underreporting issue in child mortality by adjusting the infant and under-five mortality rates in accordance with data from the Maternal and Child Health Surveillance (MCHS) system;
2) smoothed the mortality data and cause of death data with the two-dimensional mortality model and the R package “Ungroup” developed by Pascariu and colleagues.
Please see Appendix 1 for technical details and the mortality rates before and after adjustments.
> Other comments:
> 1. The discussion could be strengthened by bringing in a broader literature on age patterns of mortality and the unfolding epidemiologic transition in different parts of the world, to give the reader a better sense of how to interpret these findings and the extent to which they are unexpected. Right now the discussion is rather descriptive and the case for the importance of these results isn’t as strong as it could be.
Author response: Thank you for the advice, we have enriched the discussion by adding more information about the epidemiological transition and change of mortality patterns in other parts of the world in comparison with our findings. Please see the revised discussion section on page 11-15.
> 2. Paragraph, Line 259: It would be good to draw on the work of (Seligman, Greenberg, & Tuljapurkar, 2016), who examined this specifically.
Author response: Thank you. Indeed it’s a very nice paper, and we’ve now cited it in line 245 on page 12.
> 3. It was an interesting choice to look at cause-deleted life tables rather than a full age and cause of death decomposition, particularly when looking at the age components in the appendix. Do they differ much? What was the justification for using one strategy over another? Was this to make it easier to see the impact of each cause simultaneously on life expectancy and lifespan disparity?
Author response:
By using the cause-deleted method, we hope to examine the contribution of a cause and also build a counterfactual scenario to evaluate the potential improvement in life expectancy and lifespan disparity if a cause was removed. As shown by Beltran-Sanchez, Preston and Canudas-Romo (2008) results should not differ much between cause deleted and cause-decomposition analyses. But we were particularly interested in the interpretation of scenarios in the cause-deleted analysis, which could bring a new perspective to the lifespan inequality debate, since the causedecomposition analysis had been successfully used in the study of e-dagger (Aburto and van Raalte 2018).
References:
Beltrán-Sánchez, H., Preston, S.H. and Canudas-Romo, V., 2008. An integrated approach to cause-ofdeath analysis: cause-deleted life tables and decompositions of life expectancy. Demographic research, 19, p.1323.
Aburto, J.M. and van Raalte, A., 2018. Lifespan dispersion in times of life expectancy fluctuation: the case of Central and Eastern Europe. Demography, 55(6), pp.2071-2096.
> 4. Line 76-77: I love this article, but it is more about a new life expectancy decomposition technique than about measuring dispersion of length of life in a population (at least that’s not the main purpose of the article).
Author response: Thank you, we have now fixed the citation in line 74-76 on page 5.

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

> 5. Comments on the figures: (a) Why the reverse scale for lifespan disparity? I've never seen that and it threw me off at first, when I imagined that China was experiencing radically different patterns. (b) The legend is a little busy. Since there is a lot of white space it might be easier to make better use of direct labeling on the figure so that the reader doesn't have to look back and forth from the figure to the legend.
Author response: Thank you for the suggestions. We have revised the graphs and changed the reverse scale in all three graphs. The legend in figure 1 has been removed, and labels were added in figure 1 and 2 to make it easier for the readers to receive the information in the three figures.
>
>
> References
> Engelman, M., Canudas-Romo, V., & Agree, E. M. (2010). The Implications of Increased Survivorship for Mortality Variation in Aging Populations. Population and Development Review, 36(3), 511-539.
> Robine, J.-M. (2001). Redefining the Stages of the Epidemiological Transition by a Study of the Dispersion of Life Spans: The Case of France. Population: An English Selection, 13(1), 173-193.
> Seligman, B., Greenberg, G., & Tuljapurkar, S. (2016). Equity and length of lifespan are not the same. Proceedings of the National Academy of Sciences, 113(30), 8420-8423. doi:10.1073/pnas.1601112113
>
>
>
> Reviewer: 2
> Dr. Markus Sauerberg, Austrian Academy of Sciences Vienna Institute of Demography
> Comments to the Author:
> The study „Urban-rural lifespan disparities and cause-specific contributions: Evidence from China" includes an extensive empirical analysis, addressing urban-rural as well as male-female gaps in life expectancy and lifespan disparities, and quantifies the contribution made by causes of death. Accordingly, the research refers to several important topics, making it relevant for a broad readership. I enjoyed reading the clearly-written manuscript and appreciated all the useful information (on method, data, and figures) provided by the authors in the appendix. From my point of view, the study is almost ready for publication. My suggestion for further improvement is adding more information on the interpretation of results in the discussion section.
> Health and mortality (especially when data on specific causes of death is available) can be attributed to past health behaviours, risk, and exposures of individuals. These will likely differ between women and men as well as between urban and rural regions. On page 14 (first paragraph), the authors discuss their findings on lung and liver cancer in terms of gender differences in smoking and drinking (which is great!). I would welcome more discussion in this direction. For example, what about external causes? On page 11 (third paragraph), the authors describe that external causes have a stronger influence on life expectancy and lifespan disparity for men. Could this be connected to differences in occupational hazards (risks of accidents) between women and men (see e.g., Zhu et al. 2019 or He et al. 2005)? How do the findings relate to differences in urban and rural regions? I could imagine that pollution in urban areas might play a role (e.g., Somboonsin and Canudas-Romo 2021)? I believe that adding some information in this regard would be extremely useful to the readers.
Author response: Thank you so much for the encouraging words. The advices are very helpful and pointed the directions in which we could improve our work. Following your suggestions, we've enriched the discussion section by adding more interpretation of results. Please see page 11-15 for the revised discussion section
> I also have some minor comments and suggestions that I hope can improve the paper:
> - The definition of urban and rural (page 6, data source section): I was wondering about the classification, i.e., based on what do we distinguish between urban and rural China?
Author response: In this paper, we use the legal definition of urban and rural given by the Chinese

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

government. The urban area is defined as areas with population density higher than 1,500 persons per square kilometre, or the areas where local government of the county level or higher are located (including the adjunct areas connected by public facilities). A county-level code is given by the China National Bureau of Statistics to help with the classification. The population is readily labelled as "urban" or "rural" in the Disease Surveillance Points (DSP) system.
In the revised manuscript in line 98-99 on page 6 we have now included a sentence mentioning the definition of urban and rural China, and in Appendix 1 we also include a footnote about the classification of urban and rural areas.
> - The authors calculated life expectancy and lifespan disparity with life tables for ages between 0 and 85. I agree with the approach (given the mentioned data limitations), but I was wondering about the terminology. Wouldn't it be more appropriate to refer to truncated (or partial) life expectancy and truncated (or partial) lifespan disparity as you do on page 7 "Measure of lifespan inequality")? Please note that I would not recommend rerunning all the empirical analyses just to get the ggplots with different labels. The current labels with "(Age 0-85)" making it already very clear to the reader to which age span the measures refer.
Author response:
Thank you for the comments. It would indeed be more appropriate to refer to it as truncated life expectancy and truncated lifespan disparity throughout the paper. After the revision, we have improved our methods and use the full life expectancy and full lifespan disparity instead of the partial ones now.
In the revised paper, we
1) addressed the underreporting issue in child mortality by adjusting the infant and under-five mortality rates in accordance with data from the Maternal and Child Health Surveillance (MCHS) system;
2) instead of removing the data of people aged above 85, extrapolated the old-age mortality to age 110+ by the Kannisto model;
3) smoothed the mortality data and cause of death data with the two-dimensional mortality model and the R package "Ungroup" developed by Pascariu and colleagues. Please see appendix for technical details.
Please see page 7-8 and Appendix 1 for the methodological changes we made.
> - Speaking about terminology, I could not find the explanation for CVD in the manuscript. I guess this means cardiovascular disease? I would suggest mentioning it once in the text and using the abbreviation throughout the paper (not switching between the abbreviation and written-out version).
Author response: Sorry, CVD is the abbreviation of cardiovascular diseases. We didn't realize it was omitted in the manuscript, thank you for pointing this out. Following your advice, we now add CVD as abbreviation of cardiovascular diseases when it first appeared in line 117 on page 7 and use CVD throughout the paper.
> - Limitations of mortality data: The authors mention the possibility of underreporting. This reminded me of the discussion on quality of cause of death statistics in rural China (e.g., Wang et al. 2007), which also relates to the topic of this study. This might be another potential data restriction worthwhile for consideration.
Author response: Thank you for pointing this out. As discussed in the work by Wang and colleagues, there are concerns about the attribution of cause of death in the DSP system. We've added it in the limitation section of the manuscript. Please see line 309-310 on page 15.
> - The authors note the issue of competing risks of death for cause-deleted life table analyses. The assumption of independent causes of death is likely to be more problematic for some diseases than for others, e.g., external causes can be considered as more independent than CVD? Maybe this can be mentioned so readers are aware which conclusions should be drawn with more caution and which results are less affected by the assumption.
Author response: We totally agree with you on this. We've added this reminder in the paragraph about limitation of this work between line 314-316 on page 15.

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

> - The equation 4 on page 28 is equal to equation 2 on the previous page, no?
Author response: Yes, thank you. We've fixed the equations in appendix 3.
> - In Appendix 4, the authors provide a country-comparison, including estimates for Australia, Norway, and the US. I could imagine that these great efforts would remain unseen and underappreciated if not mentioned more directly in the main text.
Author response: Thank you for noticing this, the country-comparison in the appendix 5 is now mentioned in line 147-149 on page 8.
> - Appendix 6: The provided graphs are very helpful. I was only wondering whether the figure caption "death distribution" should be replaced with age-specific mortality rates? The plots show death rates on a logarithmic scale ("log_mx"), no?
Author response: Thank you for the suggestion. We've changed the titles of the figures for "Agespecific mortality rates of the four subpopulations in China".
> - Appendix 7: This is an interesting description (by the way, "the impact a certain cause…" the "of" seems to be missing). After reading it and looking at the e-dagger equation, I had the following question: Potential gains in life expectancy are not only affected by the number of deaths from a specific cause but also by the age at death. Averting deaths at younger ages will save a larger number of person-years lived and result in relatively large increases in remaining life expectancy. Further, lifespan disparity is affected by both, remaining life expectancy at age x and the age distribution of deaths. Will lifespan disparity necessarily decrease more when causes of deaths that mostly occur at young ages are deleted compared to deleting causes of death that occur more at older ages? Or can the effect be compensated because averting deaths at young ages leads to gains in longevity, making e-dagger larger for the remaining causes of deaths? My question relates also to the discussion on the age distribution of deaths and the threshold age (page 14). Does it mean that deleting causes of death occurring below this threshold age will always translate into life expectancy increases as well as in lifespan disparity decreases? As I stated above, this is only minor, but I got curious about the author's opinion on the relationship between gains in longevity and changes in lifespan disparity in terms of deleting specific causes of deaths at different ages.
Author response: This is a great question. When we worked on this paper, we found that lifespan disparity would decrease more when causes of deaths that predominantly occur at young ages are deleted, when compared with deleting causes of death that predominantly occur at old ages. As you wisely observed, the effect is compensated as averting deaths at young ages leads to gains in longevity, and it seems the level of this compensating effect is related with the number of death from the selected cause. But so far we only tested with the Chinese data and haven't summarised a golden rule about the relationship between the number of death, the threshold age, and age distribution of death by causes. In future, this could be something worth looking into with data from more countries and longer time period. In the revised manuscript, we've added more discussion on page 13 (between line 255 to 261) to help better explain our findings.
>
> References
> He J, Gu D, Wu X, Reynolds K, Duan X, Yao C, Wang J, Chen CS, Chen J, Wildman RP, Klag MJ, Whelton PK. Major causes of death among men and women in China. N Engl J Med. 2005 Sep 15;353(11):1124-34.
>
> Somboonsin P, Canudas-Romo V. Mortality attributable to fine particulate matter in Asia, 2000–2015: a cross-sectional cause-of-death analysis. BMJ Open 2021;11.
>
> Wang L, Yang G, Jiemin M, et al. Evaluation of the quality of cause of death statistics in rural China using verbal autopsies. J Epidemiol Community Health. 2007;61(6):519-526.
>
> Zhu, J., Cui, L., Wang, K. et al. Mortality pattern trends and disparities among Chinese from 2004 to 2016. BMC Public Health 19, 780 (2019).
>

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

>
> Reviewer: 3
> Dr. Andrea Nigri, University of Foggia
> Comments to the Author:
> This interesting manuscript analyses the urban-rural gaps in life expectancy and lifespan disparity by sex in China (2006–2016), quantifying the contributions made by causes of death.
> Using data from the Chinese Disease Surveillance Point System, the authors' findings support how rural residents are quickly catching up to their urban counterparts, shrinking the gap in both life expectancy and lifespan disparity. On the contrary, the gender gap remains large.
> The manuscript is well-written and my comments are few. To their credit, the authors were aware of possible main limitations (e.g. underreporting and unadjusted competing risks of death), acknowledging them in the manuscript.
> Author response: Thank you very much for the encouraging comments, the issues you mentioned were very important and we've fixed the problems and improved our data adjustment measures to address your concerns.
> 1. In my opinion, the abstract does not express this paper's potential. Therefore, I suggest briefly expand the abstract's Objectives section to better introduce the hypothesis (research questions or prior beliefs), and thus what the authors address through this investigation. This also applies to the Introduction section, which should be enriched with this same content. Similarly, as a service to readers, please briefly expand the Conclusions section (in the abstract) by introducing some notes about the contribution of causes of death.
Author response: Thank you for the helpful advice. Before we first submitted the manuscript, we had to delete much content in the abstract to accommodate the "Settings", "Primary outcome measures", and "Participants" sections required by the journal within the 300-word limit. We have re-written the Objectives, Results, and Conclusions section of the abstract as per your advice and expanded the last paragraph (line 82-87, page 6) of the introduction to better describe our research question.
> 2. Though life expectancy is well-known, lifespan may be less so. As such, I recommend motivating the choice of an e-dagger to measure the dispersion in life.
> Author response: Thank you, we have added the motivation of using e-dagger in line 127-134 on page 8.
> 3. Among the Strengths and limitations, the authors state that adjustment measures were taken to minimize the impact of underreporting. I think this may sound like an overstatement since the authors chose to simply remove the data of people aged over 85.
> 4. Following my previous point, China's data suffers from important issues of birth underreporting and an incorrect estimate of infant mortality. Though the literature provides some insights on this topic (Li et al., 2018; Merli & Raftery, 2000), in my understanding, the authors do not address this issue in the paper.
>
Author response: We appreciate your concerns about the impact of underreporting. To address comments 3 and 4, more adjustment measures have been taken in the revised manuscript. Now we 1) addressed the underreporting issue in child mortality by adjusting the infant and under-five mortality rates in accordance with data from the Maternal and Child Health Surveillance (MCHS) system; 2) instead of removing the data of people aged above 85, extrapolated the old-age mortality to age 110+ by the Kannisto model; 3) smoothed the mortality data and cause of death data with the two-dimensional mortality model and the R package "Ungroup" developed by Pascariu and colleagues. Please see line 108-114 on page 7 for our new adjustment measures and also Appendix 1
for data before and after adjustment.
> 5. The Discussion section provides the authors' findings, where they state, "After analysing the change of life expectancy and lifespan disparity of the subpopulations in China from 2006 to 2016, we find that the increase of life expectancy in China is accompanied by a reduction in the level of

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

dispersion in lifespan, which is consistent with what has been found in countries....”
Author response: Thank you, the sentence has been re-rewritten. Please see page 11-15 for the revised discussion section.
> Looking at Figure 1, the reader may clearly find an absence of correlation for the urban male population. This proves to be more problematic in Figure 2 with the case of CVD in urban females and the minor impact of respiratory diseases in urban males. Might this aspect be a result of the authors’ choice not to adjust competing risks of death? Can the Authors briefly argue and defend their choice?
Author response:
Thank you for noticing this. After we adjusted for the underreporting in child mortality and used the full life table instead of truncated life table for all subpopulations, the pattern for the urban male population now looks very similar to other population groups. We suspect the absence of correlation observed in the original graphs is related with the underreporting problem rather than the competing risks of death.
In our analysis we are considering all causes of death, for example deleting one and studying the remaining survival under an independent cause assumption. As such it is a competing risk model, however, a simple one compared to those more complex needed when studying diseases diagnosed individuals. Apologies for not describing it more clearly in the original manuscript. Treating causes of death as independent causes could be further refined if we had information on multiple causes of death. Nevertheless, that type of information was not available for Chinese subpopulations.
To address the concerns about the assumption of independent causes of death, we have now added in line 311-317 on page 15 a warning to readers that this assumption might be more problematic for some causes than others. For example, cancers and cardiovascular diseases might be less independent that external causes, and the conclusions should be taken with more caution.
> Furthermore, I ask the authors to do their best to improve (if possible) Figure 2, making it easier to read.
> Author response: Thank you for the suggestion, we’ve revised Figure 2 and made the scales and colors consistent with that of Figure 1. The busy legends were also fixed to make it easier to read.
> Literature
> - The estimation of death underreporting in the 2010 population census based on DCMD model life tables. Li Cheng, Mi Hong, Sun Lingxue. Population Research Vol. 42, No. 02, 99–112, March 2018.
>
> - Are births underreported in rural China? Manipulation of statistical records in response to China’s population policies. M. Giovanna Merli; Adrian E. Raftery. Demography (2000) 37 (1): 109–126.
>
> Reviewer: 1
> Competing interests of Reviewer: I have no competing interests
>
> Reviewer: 2
> Competing interests of Reviewer: For full transparency, I would like to mention that I am involved in a research project with the second author of the study.
>
> Reviewer: 3
> Competing interests of Reviewer: I declare that I have no competing interests

VERSION 2 – REVIEW

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.

REVIEWER	van Raalte, Alyson Max Planck Institute for Demographic Research, Survival and Longevity
REVIEW RETURNED	17-Nov-2021

GENERAL COMMENTS	You have done a wonderful job of addressing the reviewer concerns. I have no further concerns.

REVIEWER	Sauerberg, Markus Austrian Academy of Sciences Vienna Institute of Demography
REVIEW RETURNED	09-Nov-2021

GENERAL COMMENTS	The authors did an outstanding job in addressing my comments and I appreciate the highlighting of the corresponding passages in the revised manuscript. This makes reviewing very pleasant. Thanks! My main concerns referred to the discussion and interpretation of results. For the new manuscript, the authors made a great effort in overcoming data limitations by adjusting, smoothing, and extrapolating mortality data. This could tackle most of my previous critique. The remaining issues, which cannot be solved analytically (e.g., the assumption of independent causes of death or differences in the quality of causes of death data between urban and rural areas), are mentioned and discussed in the text appropriately. I do not have any further relevant remarks. In sum, the study provides important new results on mortality trends in China using a rich data source and a well-established and up-to-date methodology. It would love to see it published. The only minor thing which comes to my mind as an additional suggestion is including a short note on the interpretation of life expectancy and lifespan disparity. Both measures relate to the period life table population and should not be interpreted in terms of an actual life course. This has been implicitly described by the authors in the method section (the measures refer to two periods, 2006 and 2016, instead of birth cohorts). I have realized that recent articles using period indicators say this more explicitly in order to avoid potential confusion (e.g., Aburto et al. 2021 or Goldstein & Lee 2020). Again, this is only minor and not a request. References José Manuel Aburto, Jonas Schöley, Ilya Kashnitsky, Luyin Zhang, Charles Rahal, Trifon I Missov, Melinda C Mills, Jennifer B Dowd, Ridhi Kashyap, Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries, International Journal of Epidemiology, 2021 Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. Proc Natl Acad Sci U S A. 2020 Sep 8;117(36):22035-22041. doi: 10.1073/pnas.2006392117.

BMJ Open: first published as 10.1136/bmjopen-2021-050707 on 15 February 2022. Downloaded from http://bmjopen.bmj.com/ on December 5, 2024 by guest. Protected by copyright.