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)	Association of kyphotic posture with loss of independence and mortality in a community-based prospective cohort study: The Locomotive Syndrome and Health Outcomes in Aizu Cohort Study (LOHAS)
AUTHORS	Hijikata, Yasukazu; Kamitani, Tsukasa; Sekiguchi, Miho; Otani, Koji; Konno, Shin-ichi; Takegami, Misa; Fukuhara, Shunichi; Yamamoto, Yosuke

VERSION 1 – REVIEW

REVIEWER	Lorenz Uhlmann Novartis AG
REVIEW RETURNED	04-Jul-2021

GENERAL COMMENTS	General comment: Thank you for asking me to review this very interesting paper. I enjoyed reading it. The overall structure is very good. The research objective and the analysis strategy are clearly described. The results are interpreted carefully and the strengths and limitations of this study are presented and discussed. However, I have some comments which I would like to ask the authors to clarify: Major: - P. 7, line 133: You say that the primary outcome was the mortality rate. However, in your statistical analysis, you rather do a time-to-event analysis. I would recommend to adjust the wording here ("time to death" or something like that). - P. 8, lines 142-143: The categorization of BMI makes clinical sense. However, from a statistical perspective, the inclusion of the original (continuous) variable may be preferable. I think either way is acceptable. However, may I ask you if you also fit the models with BMI as a continuous covariate? Did the results change? - P. 8, lines 149-150: I guess you calculate the median and the interquartile range for WOT. This is a bit confusing, because you do categorize the WOT, if I understood you correctly. Please clarify and make sure it is getting clear in this section. - P. 8, lines 155-156: How did you check the proportional hazard assumption? Was it a test or some graphical evaluation? Or maybe some other way? Furthermore, in which sense did you do this check, meaning which covariates did you consider? - P. 8, lines 155-158: If I understood the WOT correctly, you measure it in cm on a continuous scale. After that, you categorize it into three categories. Would it make sense to include the WOT as a continuous variable? From a statistical perspective, it is usually advisable to use the original continuous variable in a model, instead of a categorized version of it. However, a categorization may also

make a lot or even more sense in specific situations, of course. Please provide some further insights.
- P. 8, lines 158-160: Why was the multiple imputation done only in a sensitivity analysis and not in the main analysis? Furthermore, did you also use the multiple imputed data to analyze your secondary endpoints? I did not see any results.
- P. 11, line 194: Please mention how exactly you calculated the mortality rate per year.
- P. 13, lines 223-226: Just for my understanding: Does that mean that you estimated two separate models (one only including female and the other only including male patients)? In an additional analysis, you could fit a model based on all patient and include an interaction term between kyphotic posture and sex. Based on that interaction term, you could give a clear statement on the difference between female and male.

Minor:
- P. 3, line 52: A hazard ratio is not bound to a specific (median) follow-up time (point). I know that this is not what you mean, but I would suggest to put these two statements (on the median follow-up time and on the aHRs) in two different sentences. You could simply write something like "The median follow-up time was 5.8 years. The aHRs for mortality…". The same applies to p. 10-11, lines 191-193.
- P.3, line 52: You provide hazard ratios for mild and severe kyphotic posture groups. Please clarify what you use as a reference group. I know that it is clearly the non-kyphotic group, but it might help the reader to mention it here. To add on my suggestion above, you could write something like "The median follow-up time was 5.8 years. Compared to the non-kyphotic group, the aHRs for mortality…".
- P. 3, line 53: You may want to introduce the abbreviation "aHR" (see also my comment later on).
- P.4, line 64: What do you mean by a "tracking ratio". If it was 98.5%, why did you only include 1,621 patients in your analyses? It becomes very clear in the text later on, but it may help to add a short explanation. Instead of using the term "tracking ratio", you could write something like "only 1.5% were lost to follow up due to…".
- P.6, lines 67-68: I think the sentence should be revised: "Failure … fails".
- P. 8, lines 150-151: You also have BMI and health status in your data. Those two variables have more than two categories (and are therefore not dichotomous). You could just say "categorical variables". Furthermore, I would prefer to say "absolute and relative frequencies", since "numbers" is too general.
- P. 8, lines 152-153: This sentence is not clear to me. What do you compare (with what)? What do you mean by "predetermined endpoint"? I guess it gets clearer in the next sentences, but I would either make this a very short sentence (stop after "employed") or be more clear in this sentence. Very minor: I would rather say "applied" instead of "employed", but this may be rather a matter of taste.
- P. 9, line 83: I guess you mean "good, very good, or excellent" health status. Please clarify and/or I would suggest to mention in the "Baseline Covariates" section that you will call this category "good" health status in the remainder of the article.
- P.9, line 184 – p. 10, line 186: I guess you meant to say high/low compared to the no kyphotic posture group, right? Please clarify. Furthermore, please define "overweight patients", for example, by adding "(>=25)", if this is what you mean. Please also refer to Table 1 at the end of this subsection.
- P. 11, line 201: My apologies if I missed it, but I think that "aHR"

was not introduced so far.
- P. 11, line 220-221 (Table 4): Please mention that you are talking about a composite endpoint. I know that you are clear on that in the text. However, looking at the table only, this does not get clear and may lead to some confusion. Please clarify in the title and also in the table itself or add some footnote.
- P. 14, lines 231-233: I appreciate that you do not overinterpret these results but just describe the results. However, I think it may still be worth mentioning that the overall number of events is very low (which makes it hard to make any conclusion).
- P. 14, line 239 – p. 15, line 240: This (the dose-response relationship) also relates to my comment above asking whether you think it may make sense to consider WOT as a continuous variable in the model.
- P. 15, line 241 and p. 16, line 284: The use of "risk of mortality" is a bit critical. I guess you are actually referring to "hazards". Please clarify.
- P. 17, line 287: My apologies, if I missed it, but how is "worst degrees of kyphosis" defined. Is it the same as your "severe" group?
- P. 18, line 317: The first sentence is a very strong statement, since it sounds like it is a final proof that kyphotic posture is (in fact) associated with LOI and mortality. However, there are some limitations (which you outlined in the strengths and limitations section). Therefore, I would suggest to weaken it a little, for example, by writing something like "This study suggests that…", or "Our data support the assumption that..." or (simply) "Kyphotic posture appears to be associated" etc.
- P. 27, Figure 2: I would suggest to cut down the y-axis to a reasonable maximum value. I do see the point why you would want to go up to 100%, but the consequence is that you can barely distinguish the three lines from each other. Furthermore, please clarify what model or test the p-value is based on (for example, in a footnote).

REVIEWER	Jorge H. Villafane
	IRCCS Fondazione Don Carlo Gnocchi
REVIEW RETURNED	20-Sep-2021

GENERAL COMMENTS	TITLE
	1. Ok
	ABSTRACT
	2. Ok
	INTRODUCTION
	3. Could you use the recent bibliography about Kyphosis (PMID: 24162521 and PMID: 29330576).
	METHODS.
	4. ok
	RESULT
	5. OK
	DISCUSSION
	6. Could you use the recent bibliography about low back pain measurement (PMID: 26732899).
	References
	7. Ok
	Tables
	8. Ok
	Figures
	9. Ok

REVIEWER	Natalie Filmann Institute of Biostatistics and Mathematical Modeling, Goethe-University, Frankfurt/Main
REVIEW RETURNED	30-Sep-2021

GENERAL COMMENTS	Hijikata and Colleagues present a thoroughly conducted manuscript adressing the association of kyphotic posture and several outcomes as mortility in a cohort of elderly japanese people. The results are presented clearly and the manuscript is very well readable. I have only a few comments regarding the statistics. 1. page 8 line 161: You wrote that regarding LOI as a secondary outcome, participants were censored after moving out of the target area, upon mortality, or on March 31, 2014. In my opinion dead should be considered as a competing risk. Could you explain to me, why you did not use competing risk analysis here? 2. Of course, kyphotic posture is correlated with age. For adjusting you used a multivariate Cox-PH model. As you have a large sample size did you also consider some kind of matching? 3. Regarding Tables 2-5, 2nd columns: percentages might be misleading as in survival analysis they should be associated with a specific point in time.

REVIEWER	Marek Zak The Jan Kochanowski University
REVIEW RETURNED	03-Oct-2021

GENERAL COMMENTS	The study was conducted in the period spanning 2008 – 2014. As 7 years have passed since, one might well assume that the outcomes could be different, should the study be carried out presently, as other factors might also come into play (e.g. as smartphones are commonplace now, their frequent use by the seniors might account for their perceptible, forward head tilt). The study fails to provide any information on whether the sternum angle and the sacral angle were examined. These two variables happen to be of particular consequence for adopting a kyphotic posture. As far as the methodology is concerned, there is no information on how the actual test was carried out, e.g. 1. Did the subjects wear their dentures or not during the procedure? When the dentures are not in place, the actual positioning of the subject's head is affected, so this factor may well prove instrumental in distorting its outcome. 2. Was the test carried out while inhaling or exhaling? Breathing naturally alters the positioning of the subject's chest. Why wasn't the test carried out with the aid of an inclinometer? This would make it objective, as well as facilitate drawing reliable conclusions. Please note that WOT is not fully objective as the method of choice, so some of the measurements may inadvertently have been taken erroneously. Also, there no information what equipment was used and how exactly the subject's handgrip strength was tested. This again is a notable methodological shortcoming.

| | The Conclusions section contains merely a reiteration of the findings already comprised in the Results section. |
| | Although the Authors boast fluent command of discursive English, their study would still appreciably benefit from a thorough flushing out by an English native speaker, so that all the rough edges might be polished off, and overall flow of the discourse itself further enhanced, as much deserved by the subject matter at issue. |

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 Mr. Lorenz Uhlmann, Novartis AG

General comment:

Thank you for asking me to review this very interesting paper. I enjoyed reading it.

The overall structure is very good. The research objective and the analysis strategy are clearly described. The results are interpreted carefully and the strengths and limitations of this study are presented and discussed. However, I have some comments which I would like to ask the authors to clarify:

<u>Response</u>

We are grateful to you for your valuable suggestions regarding our manuscript. Your constructive comments have allowed us to improve our paper significantly. We provide a response to each of your comments below.

Major:

1. **- P. 7, line 133: You say that the primary outcome was the mortality rate. However, in your statistical analysis, you rather do a time-to-event analysis. I would recommend to adjust the wording here ("time to death" or something like that).**

<u>Response</u>

As you pointed out, we have changed the "mortality rate" to "time to mortality" (line 136, page 7 of the clean copy of the revised Main Document).

2. **- P. 8, lines 142-143: The categorization of BMI makes clinical sense. However, from a statistical perspective, the inclusion of the original (continuous) variable may be preferable. I think either way is acceptable. However, may I ask you if you also fit the models with BMI as a continuous covariate? Did the results change?**

<u>Response</u>

When BMI was used as a continuous variable, the adjusted hazard ratios for mortality in the group of mild and severe kyphosis posture compared with those of the non-kyphotic posture group were 1.22 (0.73–2.03) and 2.10 (1.26–3.47), respectively, and these results were not significantly different from the results obtained using BMI as a categorical variable.

3. **- P. 8, lines 149-150: I guess you calculate the median and the interquartile range for WOT. This is a bit confusing, because you do categorize the WOT, if I understood you correctly. Please clarify and make sure it is getting clear in this section.**

<u>Response</u>

Thank you for pointing this out. We summarised the WOT as a categorical variable with real numbers and percentages, but the methods section did not include an explanation on how the categorical values were summarised. We have added "or categorical" to the sentence to make this clear (line 157, page 8 of the clean copy of the revised Main Document).

4. **- P. 8, lines 155-156: How did you check the proportional hazard assumption? Was it a test or some graphical evaluation? Or maybe some other way? Furthermore, in which sense did you do this check, meaning which covariates did you consider?**

<u>Response</u>

We performed a test using the Scaled Schoenfeld residuals and found that the proportional hazards were preserved. The variables considered in the model are listed in Table 1 (age, sex, body mass index, smoking habit, lumbar spinal stenosis, low back pain, health status, stroke history, and handgrip strength). In addition, to make it easier for the reader to understand, we have added the names of the variables in the *Methods* (lines 162-163, page 8 and line 164, page 9 of the clean copy of the revised Main Document).

5. **- P. 8, lines 155-158: If I understood the WOT correctly, you measure it in cm on a continuous scale. After that, you categorize it into three categories. Would it make sense to include the WOT as a continuous variable? From a statistical perspective, it is usually advisable to use the original continuous variable in a model, instead of a categorized version of it. However, a categorization may also make a lot or even more sense in specific situations, of course. Please provide some further insights.**

<u>Response</u>

As you point out, we collected data for the WOT as a continuous variable (in cm), and the primary analysis with WOT as a continuous variable showed an adjusted hazard ratio for mortality of 1.04 (95% CI 1.01–1.08), suggesting an association between the WOT and mortality. Given that the

association between the WOT and mortality may not be linear in the clinical sense, we followed the example of previous studies (revised refs of 13 and 21) and used the WOT as a categorical value in this study.

6. **- P. 8, lines 158-160: Why was the multiple imputation done only in a sensitivity analysis and not in the main analysis? Furthermore, did you also use the multiple imputed data to analyze your secondary endpoints? I did not see any results.**

<u>Response</u>

Thank you for your comment. As you suggested, we have added sensitivity analyses using multiple imputation for secondary endpoints, which is described in the *Statistical Analysis* section (lines 172-173, page 9 of the clean copy of the revised Main Document) and the results have been added in the *Secondary Analysis* (lines 219-221, page 12 and lines 229-230, page 13 of the clean copy of the revised Main Document) and Supplementary material (Supplementary table 2 and 3).

7. **- P. 11, line 194: Please mention how exactly you calculated the mortality rate per year.**

<u>Response</u>

We calculated the mortality rate per year by dividing the number of deaths in the None, Mild and Severe groups by the total number of person-years observed in each group.

8. **- P. 13, lines 223-226: Just for my understanding: Does that mean that you estimated two separate models (one only including female and the other only including male patients)? In an additional analysis, you could fit a model based on all patient and include an interaction term between kyphotic posture and sex. Based on that interaction term, you could give a clear statement on the difference between female and male.**

<u>Response</u>

Thank you for your constructive remarks. Multivariable analysis including interaction terms was not initially planned due to statistical power issues, and so it was not performed in this study.

Minor:

9. **- P. 3, line 52: A hazard ratio is not bound to a specific (median) follow-up time (point). I know that this is not what you mean, but I would suggest to put these two statements (on the median follow-up time and on the aHRs) in two different sentences. You could simply write something like "The median follow-up time was 5.8 years. The aHRs for mortality…".**

The same applies to p. 10-11, lines 191-193.

<u>Response</u>

We have amended these statements as per your suggestion (line 51, page 3 and lines 201, page 11 of the clean copy of the revised Main Document). Thank you very much for your suggestion.

10. - P.3, line 52: You provide hazard ratios for mild and severe kyphotic posture groups. Please clarify what you use as a reference group. I know that it is clearly the non-kyphotic group, but it might help the reader to mention it here. To add on my suggestion above, you could write something like "The median follow-up time was 5.8 years. Compared to the non-kyphotic group, the aHRs for mortality…".

<u>Response</u>

We have added "Compared to the non-kyphotic group," as per your suggestion (line 51, page 3 of the clean copy of the revised Main Document). Thank you very much for your suggestion.

11. - P. 3, line 53: You may want to introduce the abbreviation "aHR" (see also my comment later on).

<u>Response</u>

We had stated on page 3, line 48 of the clean copy of the revised main document that the adjusted hazard ratio should be abbreviated as aHR.

12. - P.4, line 64: What do you mean by a "tracking ratio". If it was 98.5%, why did you only include 1,621 patients in your analyses? It becomes very clear in the text later on, but it may help to add a short explanation. Instead of using the term "tracking ratio", you could write something like "only 1.5% were lost to follow up due to…".

<u>Response</u>

Thank you for your constructive comments. As you suggested, we changed the phrase "98.5% tracking ratio" to "1.5% lost follow up" and added that "1.5%" was "31 of the 2,193 participants included in the study" and that the reason for dropout was change of residence (lines 63-64, page 4 of the clean copy of the revised Main Document).

13. - P.6, lines 67-68: I think the sentence should be revised: "Failure … fails".

<u>Response</u>

We made a careless mistake and have corrected it. Thank you very much for pointing it out (lines 78-79, page 5 of the clean copy of the revised Main Document).

14. P. 8, lines 150-151: You also have BMI and health status in your data. Those two variables have more than two categories (and are therefore not dichotomous). You could just say "categorical variables". Furthermore, I would prefer to say "absolute and relative frequencies", since "numbers" is too general.

Response

The categorical values were not fully listed, so we have added them. Further, "numbers and percentage..." has been changed as you suggested (lines 156-157, page 8 of the clean copy of the revised Main Document). Thank you for your constructive suggestions.

15. - P. 8, lines 152-153: This sentence is not clear to me. What do you compare (with what)? What do you mean by "predetermined endpoint"? I guess it gets clearer in the next sentences, but I would either make this a very short sentence (stop after "employed") or be more clear in this sentence. Very minor: I would rather say "applied" instead of "employed", but this may be rather a matter of taste.

Response

We have changed the confusing expression "predetermined endpoint" to "date of mortality". We have also changed "employed" to "applied" as you suggested (line 158-159, page 8 of the clean copy of the revised Main Document).

16. - P. 9, line 83: I guess you mean "good, very good, or excellent" health status. Please clarify and/or I would suggest to mention in the "Baseline Covariates" section that you will call this category "good" health status in the remainder of the article.

Response

As you suggest, we have amended the *Baseline Covariates* to include an explanation of "good health status" (lines 148-149, page 8 of the clean copy of the revised Main Document). In addition, the notation in Table 1 and the footnotes in Tables 2–5 and Supplementary Table 1 have been revised as well.

17. - P.9, line 184 – p. 10, line 186: I guess you meant to say high/low compared to the no kyphotic posture group, right? Please clarify. Furthermore, please define "overweight patients", for example, by adding "(>=25)", if this is what you mean. Please also refer to

Table 1 at the end of this subsection.

<u>Response</u>

We now mention that the comparison is based on "the non-kyphotic posture group" and that overweight refers to a "body mass index ≥25" (lines 192-194, page 10 of the clean copy of the revised Main Document). Thank you very much.

18. - P. 11, line 201: My apologies if I missed it, but I think that "aHR" was not introduced so far.

<u>Response</u>

We had stated on page 11, line 208 of the clean copy of the revised main document that the adjusted hazard ratio should be abbreviated as aHR. Thank you for your concern.

19. - P. 11, line 220-221 (Table 4): Please mention that you are talking about a composite endpoint. I know that you are clear on that in the text. However, looking at the table only, this does not get clear and may lead to some confusion. Please clarify in the title and also in the table itself or add some footnote.

<u>Response</u>

To make it clearer that it is a composite outcome, we have added a footnote (Table 4, page 13 of the clean copy of the revised Main Document).

20. - P. 14, lines 231-233: I appreciate that you do not overinterpret these results but just describe the results. However, I think it may still be worth mentioning that the overall number of events is very low (which makes it hard to make any conclusion).

<u>Response</u>

We agree with your suggestion. The sentence now reads, "Although the frequencies were very low," (lines 242-243, page 15 of the clean copy of the revised Main Document).

21. - P. 14, line 239 – p. 15, line 240: This (the dose-response relationship) also relates to my comment above asking whether you think it may make sense to consider WOT as a continuous variable in the model.

<u>Response</u>

We analysed WOT as a categorical value, considering the possibility that the strength of the association may be more pronounced in the Severe group rather than a typical linear association. We

have avoided the expression "dose response relationship" and changed it to "the association was stronger in the severe kyphotic posture group" (lines 251-252, page 15 of the clean copy of the revised Main Document).

22. - P. 15, line 241 and p. 16, line 284: The use of "risk of mortality" is a bit critical. I guess you are actually referring to "hazards". Please clarify.

<u>Response</u>

We have changed risk to hazards, as you suggested (line 253, page 15 and line 294, page 17 of the clean copy of the revised Main Document).

23. - P. 17, line 287: My apologies, if I missed it, but how is "worst degrees of kyphosis" defined. Is it the same as your "severe" group?

<u>Response</u>

Thank you for pointing out the confusing wording. "Worst degrees of kyphosis" was the definition in reference 24. We have corrected the relevant sentence to read "As approximately 40% of older adults with severe kyphosis reported to have underlying OVFs, OVFs are widely thought to be a major factor contributing to the development of kyphotic posture" (lines 296-298, page 17 of the clean copy of the revised Main Document). We have also decided not to cite reference 25 in the relevant part of the text, as it would have caused more confusion.

24. - P. 18, line 317: The first sentence is a very strong statement, since it sounds like it is a final proof that kyphotic posture is (in fact) associated with LOI and mortality. However, there are some limitations (which you outlined in the strengths and limitations section). Therefore, I would suggest to weaken it a little, for example, by writing something like "This study suggests that…", or "Our data support the assumption that..." or (simply) "Kyphotic posture appears to be associated" etc.

<u>Response</u>

As you pointed out, we understood that the statement was too strong. We have weakened the wording as you suggested (line 331, page 19 of the clean copy of the revised Main Document).

25. - P. 27, Figure 2: I would suggest to cut down the y-axis to a reasonable maximum value. I do see the point why you would want to go up to 100%, but the consequence is that you can barely distinguish the three lines from each other. Furthermore, please clarify what model or test the p-value is based on (for example, in a footnote).

Thank you for your important remarks. We have redrawn Fig. 2 with the maximum value of the y-axis set to 0.4 as you suggested. We have also added "The *P*-value was calculated using log-rank test." to the *Figure legends* (lines 424, page 24 of the clean copy of the revised Main Document).

Reviewer: 2 Dr. Jorge H. Villafane

We are grateful to you for your valuable suggestions that have helped us improve our manuscript. Our responses to each of your comments are presented below.

26. TITLE

1.   Ok

ABSTRACT

2.   Ok

INTRODUCTION

3.   Could you use the recent bibliography about Kyphosis (PMID: 24162521 and PMID: 29330576).

METHODS.

4.   ok

RESULT

5.   OK

DISCUSSION

6.   Could you use the recent bibliography about low back pain

measurement (PMID: 26732899).

References

7.   Ok

Tables

8.   Ok

Figures

9.   Ok

<u>Response</u>

Thank you very much for the references. We have cited one of them (Reference 11 of the clean copy of the revised Main Document).

Reviewer: 3 Dr. Natalie Filmann

Hijikata and Colleagues present a thoroughly conducted manuscript adressing the association of kyphotic posture and several outcomes as mortality in a cohort of elderly japanese people. The results are presented clearly and the manuscript is very well readable.

I have only a few comments regarding the statistics.

<u>Response</u>

We are grateful to you for your valuable suggestions that have helped us improve our manuscript. Our responses to each of your comments are presented below.

27. 1. page 8 line 161: You wrote that regarding LOI as a secondary outcome, participants were censored after moving out of the target area, upon mortality, or on March 31, 2014. In my opinion dead should be considered as a competing risk. Could you explain to me, why you did not use competing risk analysis here?

<u>Response</u>

As you point out, mortality can be considered as a competing risk regarding LOI as an outcome. The results of the analysis using the Fine and Gray competing risks regression, which accounted for the competing risks of mortality, are shown below. The results were similar to the Cox regression with mortality censored.

	Number of participants	Frequency of loss of independence	Occurrence rate/year	Unadjusted SHR (95% CI)	Adjusted SHR (95% CI)[a]
Kyphotic posture					
None	1147	82	0.013	Ref.	Ref.

Mild	272	38	0.026	2.27 (1.56–3.30)	1.67 (1.14–2.44)
Severe	202	51	0.048	3.36 (2.37–4.78)	1.99 (1.37–2.88)

Abbreviations: SHR = subdistribution hazard ratio; CI = confidence interval.

a Fine and Gray competitive risk regression with mortality as competing risk was used to adjust for age, sex, body mass index, smoking habits, lumbar spinal stenosis, low back pain, health status, history of stroke and grip strength.

We did not include this table in the manuscript because we thought that excessive analysis might confuse the reader. Thank you very much for your valuable advice.

28. 2. Of course, kyphotic posture is correlated with age. For adjusting you used a multivariate Cox-PH model. As you have a large sample size did you also consider some kind of matching?

Response

As you point out, age is an important confounder, and matching by age is one valid option. However, there were other confounding factors to be adjusted for, and the outcome of mortality was low at 150 (6%), so we decided to perform a multivariable analysis instead of matching to maximise the power.

29. 3. Regarding Tables 2-5, 2nd columns: percentages might be misleading as in survival analysis they should be associated with a specific point in time.

Response

Thank you for pointing this out. To avoid any misunderstanding, we have removed the percentages from the second column of Tables 2–5 and Supplementary Table 1.

Reviewer: 4 Dr. Marek Zak

We are grateful to you for your valuable suggestions that have helped us improve our manuscript. Our responses to each of your comments are presented below.

30. The study was conducted in the period spanning 2008 – 2014. As 7 years have passed since, one might well assume that the outcomes could be

different, should the study be carried out presently, as other factors might also come into play (e.g. as smartphones are commonplace now, their frequent use by the seniors might account for their perceptible, forward head tilt).

<u>Response</u>

The data used in this study covers the period up to 2014, with seven more years having elapsed by now. As you pointed out, there is a possibility that confounding factors may also update over time. As mentioned in the limitations (lines 325-326, page 19 of the clean copy of the revised Main Document), it is difficult to clarify causality due to these unmeasured confounders.

31. **The study fails to provide any information on whether the sternum angle and the sacral angle were examined. These two variables happen to be of particular consequence for adopting a kyphotic posture.**

<u>Response</u>

As you pointed out, there is no information on spinopelvic parameters that can be obtained on a standing X-ray or inclinometer, including sternum and sacral angles in the present study. These parameters provide important information on how the kyphotic posture was caused. However, our focus is on the association between the resulting kyphotic posture and mortality, and not on the cause of the kyphotic posture. We therefore believe that the lack of these information does not introduce a serious bias into the present study. We have added a *Limitation* as follows (lines 321-323, page 18 and line 324, page 19 of the clean copy of the revised Main Document).

> "Furthermore, we did not use X-rays or inclinometer to assess kyphotic posture, and so it was not possible to determine the cause of the posture. However, we believe that the absence of spinal parameters such as kyphotic angle does not introduce a serious bias, as our focus is on the resulting kyphosis posture, not on its cause."

32. **As far as the methodology is concerned, there is no information on how the actual test was carried out, e.g. 1. Did the subjects wear their dentures or not during the procedure? When the dentures are not in place, the actual positioning of the subject's head is affected, so this factor may well prove instrumental in distorting its outcome. 2. Was the test carried out while inhaling or exhaling? Breathing naturally alters the positioning of the subject's chest.**

<u>Response</u>

Unfortunately, there is no information on whether the participants wore dentures or not. As the WOT was measured at each local government health check, we assume that participants who wore dentures daily were measured with them, while those who wore them only for meals were measured without them. We also assume that there was no specification of breathing during the WOT measurement. Therefore, measurement error was inevitable in the WOT. However, since these measurement errors occurred equally in the None, Mild, and Severe groups, which would have diminished the association, we do not expect them to have a significant impact on the study results.

33. Why wasn't the test carried out with the aid of an inclinometer? This would make it objective, as well as facilitate drawing reliable conclusions.

<u>Response</u>

Thank you very much for your valuable remarks. It is true that the use of an inclinometer would have allowed us to measure kyphosis with greater accuracy than WOT. However, the WOT is much simpler and less laborious than the inclinometer and can be performed by anyone without any special training and without using any special tools. We have chosen to use the WOT because of its simplicity.

34. Please note that WOT is not fully objective as the method of choice, so some of the measurements may inadvertently have been taken erroneously.

<u>Response</u>

As you point out, WOT is a semi-quantitative measurement method with limited precision, and we have added the issue of measurement to the *Strength and Limitations* subsection as follows (line 316, page 18 of the clean copy of the revised Main Document).

"the measurement of kyphotic posture may not be sufficiently precise."

35. Also, there no information what equipment was used and how exactly the subject's handgrip strength was tested. This again is a notable methodological shortcoming.

<u>Response</u>

The equipment used to measure handgrip strength has been added to the *Baseline Covariates* (lines 151-152, page 8 of the clean copy of the revised Main Document).

36. The Conclusions section contains merely a reiteration of the findings already comprised in the Results section.

In response to your suggestion, we have revised the wording of the *Conclusion* (line 331, page 19 of the clean copy of the revised Main Document).

37. Although the Authors boast fluent command of discursive English, their study would still appreciably benefit from a thorough flushing out by an English native speaker, so that all the rough edges might be polished off, and overall flow of the discourse itself further enhanced, as much deserved by the subject matter at issue.

Response

Thank you for your comment. This manuscript has undergone multiple rounds of editing in English by Editage (https://www.editage.jp/services/english-editing). We have passed on your suggestions to Editage and the article underwent another round of proofreading.

VERSION 2 – REVIEW

REVIEWER	Lorenz Uhlmann Novartis AG
REVIEW RETURNED	23-Dec-2021

GENERAL COMMENTS	I would like to thank the authors for addressing all my comments carefully. I am satisfied with the replies and the corresponding updates done in the manuscript. I have only a few additional minor comments: - p. 9, line 150: The word "for" is used twice. I guess the first one needs to be deleted. - Throughout the document: I am more used to read "p-values" instead of "P-values". I would suggest to use a lower "p".

REVIEWER	Natalie Filmann Institute of Biostatistics and Mathematical Modeling, Goethe-University, Frankfurt/Main
REVIEW RETURNED	09-Dec-2021

GENERAL COMMENTS	I have no further comments. Thank you for considering my suggestions.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 3   Dr. Natalie Filmann

I have no further comments. Thank you for considering my suggestions.

Response

We sincerely appreciate your constructive comments.

Reviewer: 1   Dr. Lorenz Uhlmann

I would like to thank the authors for addressing all my comments carefully. I am satisfied with the replies and the corresponding updates done in the manuscript. I have only a few additional minor comments:

- p. 9, line 150: The word "for" is used twice. I guess the first one needs to be deleted.

Response

Thank you for pointing out the grammatical error; we have corrected the relevant portion as follows (line 150, page 8 of the clean copy of the revised Main Document):

Before correction: "…a validated diagnostic support tool for specifically designed for this purpose".

After correction: "…a validated diagnostic support tool specifically designed for this purpose".

- Throughout the document: I am more used to read "p-values" instead of "P-values". I would suggest using a lower "p".

Response

Thank you very much for pointing out the details. We have changed "P" in the text to "p" as you suggested (line 206, page 11 and line 430, page 24 of the clean copy of the revised Main Document, and Figure 2).