

# **A simulation study of the impact of population-wide lifestyle modifications on life expectancy in the Chinese population**

Corresponding Author: Professor Jun Lv

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The present anaalysis was based on big databases and well done. The manuscript was generally well written. There are some suggestions for revision of the manuscript.

1. A major concern is the lack of novelty. The authors proved the truth is true, "if one lives healthier, he or she will live longer". In fact, the real problem in regard to lifestyle modifications is how to achieve a healthier lifestyle and what the bearriers are. The authors may consider to utilize these databases to do some useful analysis.
2. Also in line with the above concern, the study lacks of clinical significance. Even though the authors did not do useful analysis, they may need to provide some clear cut suggestions on the trategies for achieving healthier lifestyle and let the policy makers know what they can and should do.
3. Though the message is very simple, the analysis is complicated for the average readers of the Journal. The manuscript may be more suitable for a specialty epidemiology journal.
4. Table 2 is big. Are all those listed categories used in the analysis or useful ?
5. Liftstyle is important for healh. However, healthcare is probably more important. In older days, for instance, more than a thousand years ago, people did not smoke cigarettes, did not drink spirits, did have a lot of exercise, and ate fresh vegetables, but they did not live much longer. The authors may need to take this into account.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

Thank you for the opportunity to review "The impact of population-wide lifestyle modifications on improving life expectancy in the Chinese population: a simulation study." I am the lead author of the "Canadian" study referenced in the introduction.

Overall, this strong study was presented clearly. That said, I had difficulty assessing parts of the study, and some of my comments may be incorrect because of the assumptions I made to fill missing gaps. Based on the authors' responses, I would be pleased to review the manuscript again.

For example, Supplementary Table 4 provide hazard ratios for the risks. However, I could not find information on how the risks were adjusted (unadjusted, age and sex, full-adjustment?). An HR of heavy smoking of 1.3 for men and 1.6 for women seems unusually low. A more typical HR greater than 2 will considerably affect the studies' conclusions. The low HR likely reflects a highly specified model with many correlated risks. However, it is difficult to interpret these low smoking HR findings without knowing which risks were included in the model.

Strengths of the study include:

- 1) A large national study that included well-measured exposures (for a population health survey) for sociodemographics, health behaviours, biophysical measures, and health status measures (i.e. diseases) that are leading risks for all-cause mortality
- 2) The use of a multivariable risk approach. This is a more sophisticated approach than other common approaches with strengths compared to others.
- 3) The study used a well-established approach to developing a predictive algorithm - with caveats (see below). Prespecification for many parts of the study was a particular strength.
- 4) The China study population was welcomed. I have not read a similar study performed in a Chinese population. Having a broader range of countries helps advance knowledge.

Notable findings and significance.

The study has two noteworthy findings.

- The estimates of risk factor burden are notable and will inform health policy. These findings are informative for both China and other countries, including countries that are similar or differ in their sociodemographic and health behavioural risk.
- Developing a multivariable all-cause mortality risk prediction algorithm and its application. The number of studies and countries developing these risk algorithms is still limited. This study demonstrates that all-cause mortality can be discriminately predicted with good calibration in China, which supports similar findings in other countries—such as our study in Canada. Multivariable risk algorithms have applications beyond this current study. The algorithm's good performance paves the way for additional studies and applications.

Major comments:

1) Report on TRIPOD-AI, CHEERS and STROBE or GATHER reporting guidelines. The study appears to conform well to the appropriate EQUATOR guidelines, noting a lack of reporting guidelines for noncommunicable disease simulation reporting guidelines (disclosure: I am leading an international consensus study to develop these guidelines.) The study prominently includes a predictive algorithm -- TRIPOD may be appropriate. Journals sometimes use CHEERS for simulation modelling studies. GATHER has been used for burden studies. Some of my comments below reflect potential gaps in reporting for those guidelines.

2) Title: Can a more specific study type be included? Perhaps a "simulation model" rather than a "simulation study." It can be very difficult to identify these studies in systematic reviews. To ensure your study appears in reviews, use terms such as "simulation" and "model." MESH or keywords can be helpful: "Risk factor burden model," "Policy option or scenario model," etc.

3) The predictive variables were helpfully described, and the diet exposures were extensively discussed. However, I had difficulty following the final form of the models.

Provide additional details on the model specification(s). Provide a table of exposures and their form for the three models (full model, reduced model and model supporting Supplementary Table 4). This could include the variables, their form (i.e. number of degrees of freedom, spline knots, etc., max/min and definition)

4) Describe any tests of assumptions for the Cox model. Was the proportionality of the hazards assessed?

5) Provide the final model with beta coefficients and an example calculation -- as recommended by TRIPOD AI.

6) The burden results are smaller than expected. Can you comment or provide more details?

7) The smaller-than-expected burden estimates may be from the highly specified model, including mediators and correlated variables with health behaviours. Exposures such as blood pressure and diseases are mediators ("downstream") from health behaviours. People report values in measures such as self-rated health that consider their health behaviours in their reports. Sociodemographic measures such as marital status may also be correlated with health behaviours.

Can you discuss the sources of measurement error, bias, and model specification for your model and its findings further? Ideally, the discussion should include the direction and potential magnitude of the effect. It is common practice to perform analytic assessments using various approaches.

Building on #7.

8) It may be helpful to add a correlation matrix of variables.

9) Describe the influence of predictors on your final model. This could provide the predictors' HR (and beta coefficients). However, these can be difficult to interpret for splines and interactions; other methods could be used.

10) Consider assessing calibration across subgroups - particularly groups with potential interpretation bias, such as sociodemographic groups. See recommendations in TRIPOD-AI.

11) It is increasingly recommended that continuous calibration plots be used, but this is a minor point that does not require revision unless you are planning additional analyses.

12) Bootstrap validation is also increasingly recommended, but this is a minor point that does not require revision unless you

are planning additional analyses.

13) Further, backward selection has noted limitations.

14) In line 252, it seems that the equation in the math display.

The terms  $S_0(5)$  and  $\exp(V)$  should be multiplied, whereas in the paper,  $S_0(5)$  is raised to the power of  $\exp(V)$ .

(Remarks on code availability)

Including the code for the model, development would be helpful, but this is still not commonly provided in papers. (I strive to publish code for our new studies.

Reviewer #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

This is a good and important paper. It is for me however as a trained mathematician, not an easy paper to read. Below I state some minor recommendations which might improve the accessibility of the manuscript. I also raise some methodological queries. First I will provide an overall summary of my thoughts.

The manuscript produces novel and important results pertaining to the impact of lifestyle changes on the life-expectancy of Chinese population members. These results are of great policy interest. The authors claim that their model provides a "new analytical strategy". I do not fully agree with this and believe that their model is a nuanced application of previous work to a specific case study of China (see below). Overall I was happy with the methodological approach, although I did have a slight reservation regarding the use of Cox Proportional Hazards models (see below). Clarification either of my misunderstanding of this approach and as to why the choice of model is appropriate, or clarification of the limitations of the approach would however suffice to ease my issues. The detail of the results is a strong aspect of this work and I am more than happy with how the results are presented. Overall, after addressing my minor concerns I would be happy for this work to be published.

Minor issues:

To begin with, after reading the paper and its emphasis on the effects that lifestyle habits have on life expectancy, I am not sure whether the authors refer to period based life expectancy or cohort based life expectancy. Are the authors assuming a stationary population with non-time-varying death rates? This should be clarified very early on as the two interpretations of life expectancy are very different, and have different uses for policy.

I am not fully clear how this model differs/extends what has already done in the referenced model of Manuel et al (2016). You claim that your model is an analytical innovation, but I do not see how. I do believe however that your work is important as an applied analysis specific to the Chinese population. Some other points here: you criticise the "traditional" aggregated data approach for not combining interrelating predictors, but then criticise the multivariable approach for not considering effects separately. I believe that your model's strength is in the latter approach, but I think it is unfair to be critical of Manuel et al (2016) when your paper relies heavily on that approach. Moreover, the authors should state that they apply the same methods as Manuel et al (2016), i.e., Cox Proportional Hazards to model mortality, and use a similar multivariable approach.

On page 8 you explain how Cox Proportional Hazard functions are used to calculate 5-year probabilities of death. I am not convinced that the assumption of constant hazard ratios over time is appropriate. The effects that predictor variables have on mortality risk may vary over the follow up period. It is also unclear to me whether age is a time-variant variable. Not considering (multiple) predictor variables as time-variant might mis-estimate the true effects of any one of them. For instance Body Mass Index (BMI) and blood pressure are likely to change over the follow up period and be correlated with mortality. I would like at least a discussion on this, or the authors to clarify how this concern of mine is not a problem in their model.

On page 10, you go through some results, and refer to Figure 2 in the supplement. I found it very hard to interpret the figure, and it did not guide me towards the significance of your findings. In particular, could you provide more detail as to what the ROC curve means/how to interpret it? Personally, I have not heard of a "c-curve", nor AUC or ROC. A simple explanation will make the paper more accessible to a wider audience.

In the results, I am not sure what sources of uncertainty the confidence interval cover. Are possible errors in sampling covered here?

Grammar, use of language and other:

The first sentence need splitting into at least two. As it stands it is a poor introduction to an interesting paper.

The second paragraph states "Healthy SG" without stating what the acronym stands for.

The manuscript could do with a re-reading and checking of all grammar.

(Remarks on code availability)

No code supplied

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

No further comments.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I commend the authors for their thorough and thoughtful responses to the reviewer comments. The revisions are comprehensive, and I am particularly grateful for the extensive additional analyses, including the hazard ratio correlation matrices, the assessment of potential model over-specification, and the evaluation of calibration across subgroups. These enhancements have materially strengthened the manuscript.

One comment I would add concerns the relatively low hazard ratios reported for smoking in the Chinese population. As the authors note, this is consistent with previous studies. I would suggest further acknowledging that hazard ratios are inherently relative and depend on the underlying baseline risk. In settings like China—where baseline mortality risk may be elevated due to communicable diseases, injuries, or other competing causes—the relative contribution of specific risk factors such as smoking may appear attenuated in all-cause mortality models. This does not necessarily imply a lower absolute risk from smoking but rather reflects the complexity of mortality attribution in diverse epidemiological contexts. This may also explain why smoking hazard ratios are more pronounced in disease-specific models than in all-cause mortality frameworks.

Overall, I am very satisfied with the authors' responses and revisions. I support the manuscript's publication in its revised form.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

I have read with great interest the revised version of the article "The impact of population-wide lifestyle modifications on improving life expectancy in the Chinese population".

First I would like to thank the authors for investing so much energy into revising the manuscript. The article is much improved. The research is well presented and rigorously detailed. The authors have considerably improved the clarity of the study. I am now happy to recommend that the article is published.

I would like to thank the authors for addressing my previous comments:

It is important to clarify the distinction between period LE and cohort LE. The authors have now explained this. This helps explain what their results could mean for policy decision.

This research is interesting and the fact that the methods are not novel should not in any way inhibit its publication. The authors have now clarified that their model is adopted from an established framework. That being said, the results are new, and just as important. As the research heavily relies on the framework of the Manuel et al (2016) -- and to my understanding does not advance their methods in any way -- I welcome the author's removal of criticisms [of Manuel et al (2016)].

Demonstrating that the proportional hazards assumption is satisfied (log-log; discrete and Schoenfeld residual; continuous predictors) is also important, even as an Supplement. This assumption is crucial to the methods, and should not have been overlooked in the previous manuscript. I am happy with this inclusion.

I do appreciate that word restrictions constrain authors to prioritise text. However, Nature Comms is a multi-disciplinary journal, and readers (like myself) will not be versed in the terminology of epidemiology. I welcome the explanations of the ROC curve and AUC in the Supplements.

(Remarks on code availability)

I apologise that I am not able to review the code: I am not trained in using the software STATA.

**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

## Response to the reviewers

Manuscript reference number: NCOMMS-25-02875

Title: The impact of population-wide lifestyle modifications on improving life expectancy in the Chinese population: a simulation study

*Note: The page number refers to that in the clean version.*

### **Reviewer #1 (Remarks to the Author):**

The present anaalysis was based on big databases and well done. The manuscript was generally well written. There are some suggestions for revision of the manuscript.

**Response:** We thank the reviewer for this positive comment. We have addressed the reviewer's concern and specific suggestions point-by-point in the following section.

1. A major concern is the lack of novelty. The authors proved the truth is true, "if one lives healthier, he or she will live longer". In fact, the real problem in regard to lifestyle modifications is how to achieve a healthier lifestyle and what the bearriers are. The authors may consider to utilize these databases to do some useful analysis.

**Response:** We agree with the reviewer that identifying effective measures to promote healthy lifestyles is a critical public health priority. However, this study also addresses a critical and yet unresolved problem. In 2016, the Chinese government put forward "Healthy China Initiative 2019-2030", outlining specific goals to improve the lifestyle of its population. Nevertheless, there is insufficient population-based evidence from China to verify whether the proposed one-year increase in life expectancy (LE) could be achieved if these lifestyle improvement targets were met. Moreover, whether the currently established targets need to be adjusted also warrants scientific investigation. These are all important strategic problems that demand urgent attention. Hence, this study aims to estimate the gains in life years under various hypothetical lifestyle intervention scenarios, thereby providing valuable scientific evidence to inform and optimize future policy-making for healthy lifestyle promotion.

This study has generated several noteworthy findings. For example, compared to the initiative's uniform target of reducing overall smoking prevalence to 20%, implementing a sex-specific tobacco control strategy, i.e., a reduction in male smoking rate to 20%, would result in a greater LE improvement. Notably, adopting the "tobacco-free generation" strategy, i.e., preventing adolescents from taking their first puff, could increase LE by 1.02 years. For physical activity, the population-wide strategy proved more effective than the current high-risk group intervention approach. Furthermore, we identified that concurrent increases in fresh fruits and fish or seafood consumption

would produce substantial LE gains. This finding highlights an evidence gap in current dietary guidelines, which provide no explicit intake recommendations for these two food groups.

In summary, the estimates of risk factor burden in this study will inform health policy for China and other countries, including those with similar or distinct socio-demographic and behavioral risk profiles.

2. Also in line with the above concern, the study lacks of clinical significance. Even though the authors did not do useful analysis, they may need to provide some clear cut suggestions on the strategies for achieving healthier lifestyle and let the policy makers know what they can and should do.

**Response:** With simulation analysis for each lifestyle factor, this study provides policymakers with evidence-based strategies to optimize current policies and accelerate progress toward Healthy China 2030 objectives. Specifically, our findings suggest: (1) implementing sex-specific smoking reduction targets coupled with more stringent tobacco control measures; (2) strengthening existing alcohol restriction policies; (3) adopting population-wide strategy to improve physical activity level; (4) refining the food groups in the current dietary guidelines and integrating other more healthful foods. Please refer to our response to the aforementioned comment.

3. Though the message is very simple, the analysis is complicated for the average readers of the Journal. The manuscript may be more suitable for a specialty epidemiology journal.

**Response:** We used an analytical strategy adapted from a landmark Canadian study in this field.<sup>1</sup> Compared with traditional methods used in burden studies of health-behaviour, this newly developed method is more flexible, allowing us to assess the combined effect of simultaneous changes in multiple lifestyle factors on LE of the whole population. We can therefore prospectively evaluate the potential gains in LE under various intervention scenarios. To address the reviewer's concern and improve the readability of this analysis, we provided a full description of the key steps of this method in the "Methods" section (page 16, lines 488-497).

As outlined on the official website, *Nature Communications* is a multidisciplinary journal, dedicated to publishing high-quality research in all areas, including health science. Notably, *Nature* has launched a collection titled "Progress towards the Sustainable Development Goals", aiming to assess interventions that accelerated the achievement of the Sustainable Development Goals (SDGs). Our study systematically evaluated the impact of various lifestyle intervention strategies on LE within the Chinese population, providing evidence-based policy optimization pathways that aligns with the SDGs framework.

In summary, our study employed a suitable method to address a critical policy problem, yielding clear and actionable insights. We are therefore confident that our research would be of great interest to *Nature Communications*, as well as its broad readership.

4. Table 2 is big. Are all those listed categories used in the analysis or useful ?

**Response:** Yes, the variables listed in Table 2 were the predictors in the 5-year mortality risk prediction model, which we developed for the calculation of LE under various scenarios. The formation of each variable was determined based on the Bayesian information criteria. Details of the development of the prediction model were described in the “Statistical analysis” section (page 14, lines 428-431; page 15; page 16, lines 467-477).

5. Lifestyle is important for health. However, healthcare is probably more important. In older days, for instance, more than a thousand years ago, people did not smoke cigarettes, did not drink spirits, did have a lot of exercise, and ate fresh vegetables, but they did not live much longer. The authors may need to take this into account.

**Response:** We thank the reviewer for this suggestion. However, this study emphasizes the superior importance of lifestyle factors over healthcare in reducing disease burden. In most developing countries, including China, limited economic resources necessitate primary prevention through lifestyle interventions. Without such preventive measures, the vast population base could lead to a surge in disease prevalence, overwhelming healthcare systems and further straining economic development.

Notably, while developed nations possess advanced healthcare infrastructure, their recent gains in LE primarily reflect extended years lived with chronic diseases rather than improved life quality.<sup>2</sup> These findings collectively underscore the imperative for global prioritization of lifestyle interventions as the cornerstone of disease burden reduction.



**Reviewer #2 (Remarks to the Author):**

Thank you for the opportunity to review "The impact of population-wide lifestyle modifications on improving life expectancy in the Chinese population: a simulation study." I am the lead author of the "Canadian" study referenced in the introduction.

Overall, this strong study was presented clearly. That said, I had difficulty assessing parts of the study, and some of my comments may be incorrect because of the assumptions I made to fill missing gaps. Based on the authors' responses, I would be pleased to review the manuscript again.

**Response:** We sincerely appreciate the reviewer's valuable time in reviewing our manuscript and for the positive comments. We apologize for the lack of clarity in parts of the original manuscript. We have addressed the reviewer's concern and specific suggestions point-by-point in the following section.

For example, Supplementary Table 4 provide hazard ratios for the risks. However, I could not find information on how the risks were adjusted (unadjusted, age and sex, full-adjustment?). An HR of heavy smoking of 1.3 for men and 1.6 for women seems unusually low. A more typical HR greater than 2 will considerably affect the studies' conclusions. The low HR likely reflects a highly specified model with many correlated risks. However, it is difficult to interpret these low smoking HR findings without knowing which risks were included in the model.

**Response:** We thank the reviewer for the careful review, and we are sorry we didn't make it clear. In Supplementary Table 4 (Supplementary Table 7 in the revised supplementary file), all predictor variables were included simultaneously in the model, that is, the risks were fully adjusted. We have added this explanation to the footnote of Supplementary Table 7 in the revised supplementary file and Table 2 in the revised main text to clarify this point.

As the reviewer pointed out, the hazard ratios (HRs) of all-cause mortality that we observed for heavy smokers ( $\geq 20$  cigarettes per day) relative to never smokers in this study were much lower than those reported in the Canadian population, in which the HRs of heavy smoking were 2.83 for men and 3.26 for women. This finding was consistent with what was reported in our previous publication<sup>3</sup> and other Chinese populations.<sup>4</sup> In fact, the impact of smoking on all-cause mortality risk and LE in our population was smaller than in most high-income countries. We think that in developing countries such as China, mortality risk may be influenced by a broader spectrum of risk factors, including potential environmental hazards in the home, work, and broader outdoor environment, like ambient air pollution, chemical contamination of food and water.<sup>5</sup> Thus, the relative impact of lifestyle alone might be slightly diminished in developing countries.

In addition, we also evaluated whether this modest effect size might result from a highly specified model by plotting a complete correlation matrix encompassing all predictors. The analysis revealed minimal pairwise correlations among predictors, indicating that coefficient estimates were unlikely to be substantially affected by collinearity. Please refer to our response to Comment #7 for more details.

Strengths of the study include:

- 1) A large national study that included well-measured exposures (for a population health survey) for sociodemographics, health behaviours, biophysical measures, and health status measures (i.e. diseases) that are leading risks for all-cause mortality
- 2) The use of a multivariable risk approach. This is a more sophisticated approach than other common approaches with strengths compared to others.
- 3) The study used a well-established approach to developing a predictive algorithm - with caveats (see below). Prespecification for many parts of the study was a particular strength.
- 4) The China study population was welcomed. I have not read a similar study performed in a Chinese population. Having a broader range of countries helps advance knowledge.

Notable findings and significance.

The study has two noteworthy findings.

- The estimates of risk factor burden are notable and will inform health policy. These findings are informative for both China and other countries, including countries that are similar or differ in their sociodemographic and health behavioural risk.
- Developing a multivariable all-cause mortality risk prediction algorithm and its application. The number of studies and countries developing these risk algorithms is still limited. This study demonstrates that all-cause mortality can be discriminately predicted with good calibration in China, which supports similar findings in other countries—such as our study in Canada. Multivariable risk algorithms have applications beyond this current study. The algorithm's good performance paves the way for additional studies and applications.

**Response:** We sincerely appreciate the reviewer's careful reading and recognition of the contributions made by our work.

Major comments:

- 1) Report on TRIPOD-AI, CHEERS and STROBE or GATHER reporting guidelines. The study appears to conform well to the appropriate EQUATOR guidelines, noting a lack of reporting guidelines for noncommunicable disease simulation reporting guidelines (disclosure: I am leading an international consensus study to develop these guidelines.) The study prominently includes a predictive algorithm -- TRIPOD may be

appropriate. Journals sometimes use CHEERS for simulation modelling studies. GATHER has been used for burden studies. Some of my comments below reflect potential gaps in reporting for those guidelines.

**Response:** We thank the reviewer for pointing this out. We have revised the manuscript accordingly. Please refer to our detailed responses to the comments below.

2) Title: Can a more specific study type be included? Perhaps a "simulation model" rather than a "simulation study." It can be very difficult to identify these studies in systematic reviews. To ensure your study appears in reviews, use terms such as "simulation" and "model." MESH or keywords can be helpful: "Risk factor burden model," "Policy option or scenario model," etc.

**Response:** We thank the reviewer for this suggestion. However, *Nature Communications'* formatting guidelines prohibit punctuation or puns in the title. We have to remove all subheading text that originally appeared after colons.

3) The predictive variables were helpfully described, and the diet exposures were extensively discussed. However, I had difficulty following the final form of the models. Provide additional details on the model specification(s). Provide a table of exposures and their form for the three models (full model, reduced model and model supporting Supplementary Table 4). This could include the variables, their form (i.e. number of degrees of freedom, spline knots, etc., max/min and definition)

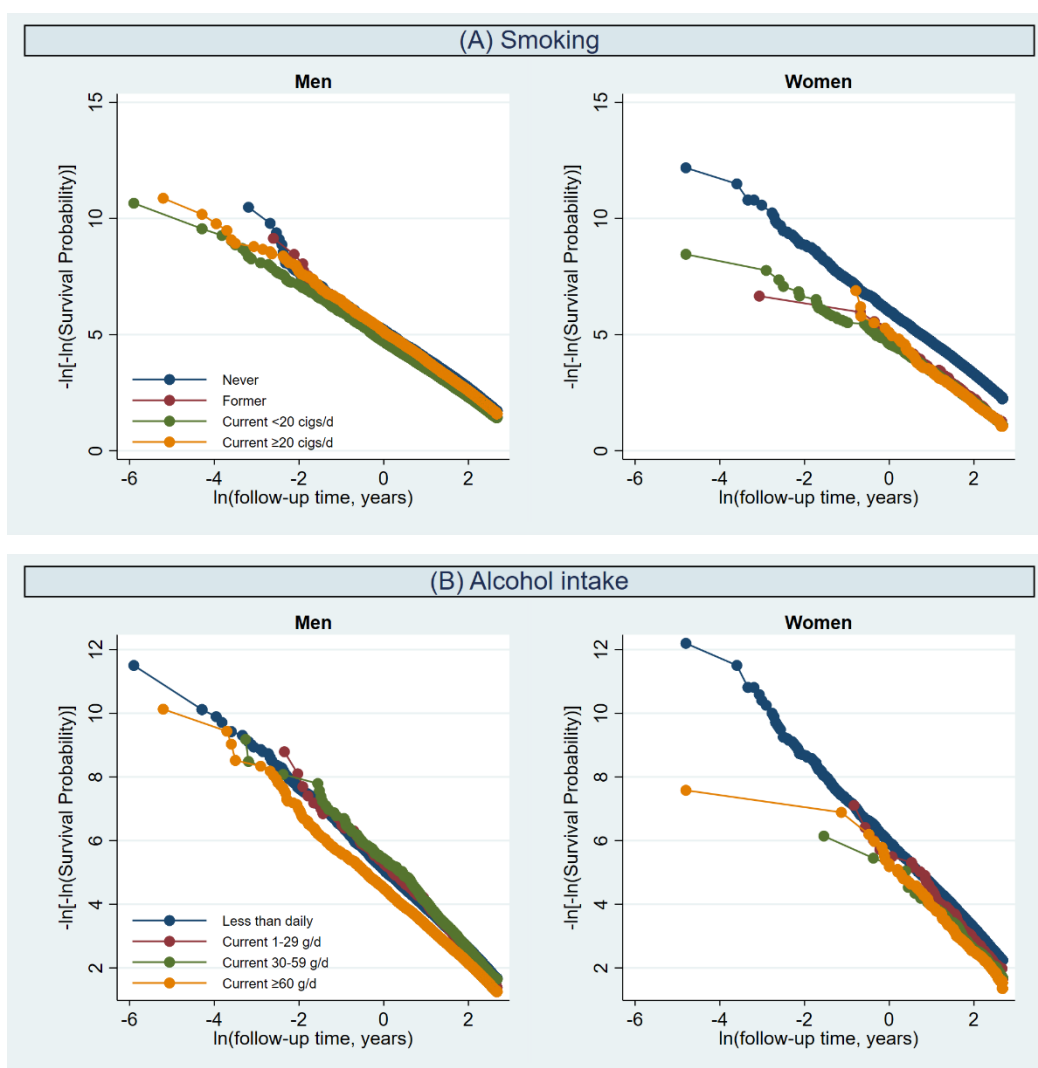
**Response:** We did not use a reduced model in this analysis, but only developed a full model to predict the 5-year all-cause mortality risk. In addition, Supplementary Table 4 (Supplementary Table 7 in the revised supplementary file) presents the variables and their coefficients included in the models constructed based on the whole male and female population, which was a validation of the prediction model constructed using a simple cross-validation approach in the primary analysis. The relevant methodological and results descriptions were given in lines 470-472 on page 16 and lines 150-153 on page 6 of the revised main text.

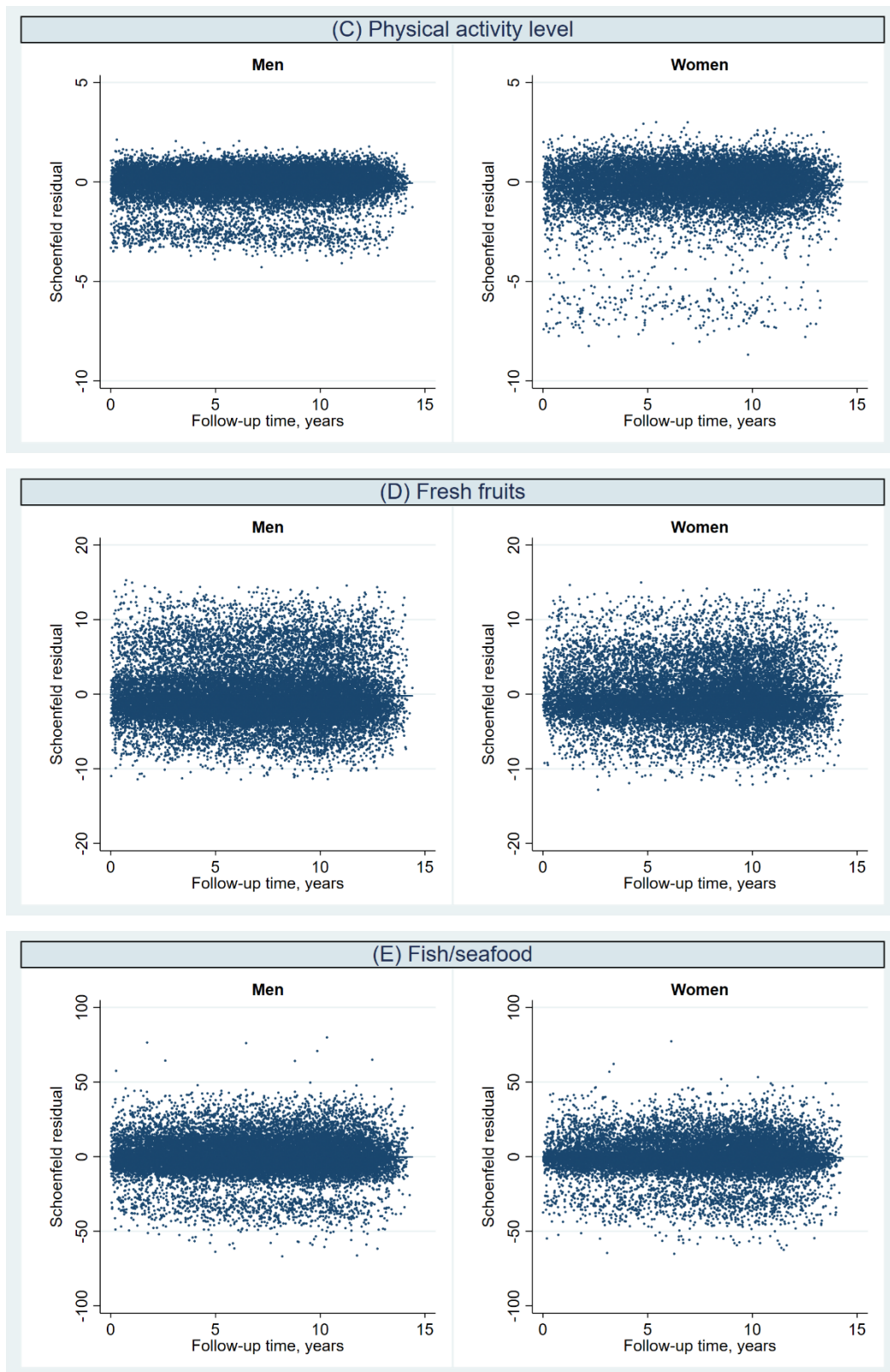
As suggested by the reviewer, we provided detailed information on the predictor variables in the full model in Supplementary Table 4 of the revised supplementary file, including the variable descriptions, types, and either the number of levels for categorical variables or value ranges for continuous variables.

4) Describe any tests of assumptions for the Cox model. Was the proportionality of the hazards assessed?

**Response:** We had performed the proportional hazards assumption test during model development, but we did not present the results, with reference to most predictive modeling research. However, to address the reviewer's concern, we present here the

test results. We assessed the proportional hazards assumption using log-log survival plots for categorical variables and Schoenfeld residual tests for continuous variables, and no significant violation was observed. The results of the five lifestyle factors used for subsequent simulation analysis are presented in the following **Figure 1**. For smoking and alcohol intake, the log-log survival plots for each category are parallel, except for a slight crossover at the very beginning of the curve for women, possibly due to the small number of current smokers (2.9%) and drinkers (1.8%). Meanwhile, the Schoenfeld residual plots of physical activity level and daily intake of fresh fruits and fish/seafood are independent of time as well, and had *P*-values greater than 0.05. This explanation was also added in the revised main text (page 6, lines 146-148).





**Figure 1.** Proportional hazards test for the five lifestyle factors in simulation analysis. (A-B) log-log survival functions for smoking and alcohol intake; (C-E) Schoenfeld residual plots for physical activity level and daily intake of fresh fruits and fish/seafood.

5) Provide the final model with beta coefficients and an example calculation -- as recommended by TRIPOD AI.

**Response:** We thank the reviewer for this suggestion. We have provided the complete risk calculation steps in **Supplementary Table 5** and the beta coefficients for all parameters in **Supplementary Table 6** of the revised supplementary file.

6) The burden results are smaller than expected. Can you comment or provide more details?

**Response:** We thank the reviewer for this insightful comment. Following the reviewer's recommendations in the next comment, we have comprehensively analyzed the possible reasons for the relatively low disease burden estimates from various aspects, such as a highly specified model and the measurement errors in population surveys. We have revised accordingly in the revised main text. Please see our detailed response to the following comment for complete discussion.

7) The smaller-than-expected burden estimates may be from the highly specified model, including mediators and correlated variables with health behaviours. Exposures such as blood pressure and diseases are mediators ("downstream") from health behaviours. People report values in measures such as self-rated health that consider their health behaviours in their reports. Sociodemographic measures such as marital status may also be correlated with health behaviours.

Can you discuss the sources of measurement error, bias, and model specification for your model and its findings further? Ideally, the discussion should include the direction and potential magnitude of the effect. It is common practice to perform analytic assessments using various approaches.

**Response:** We thank the reviewer for this insightful comment. Following the reviewer's suggestion in the next comment, we constructed a correlation matrix for all predictor variables in the prediction model to examine the inter-variable correlations. The results demonstrated only weak correlations among variables, indicating minimal impact of collinearity on parameter estimation. The correlation matrix plot was added to the revised supplementary file as **Supplementary Figure 2**, and the same description was added to the "Results" section of the revised main text (page 6, lines 144-146).

We further examined whether the lifestyle effects were mediated by other variables by comparing three progressively adjusted models: (1) age and lifestyle factors only; (2) adding sociodemographic factors, including education level and marital status; (3) further adjusting for health indicators comprising self-rated health, baseline stroke, cancer, COPD, diabetes, BMI, systolic blood pressure, and resting heart rate. The effect sizes of lifestyle factors and other exposures remained almost consistent in these three models, except that only the effects of smoking and alcohol use were slightly attenuated

in the fully adjusted model for men. This supported that the prediction model has appropriate specifications, and the effect size estimates were robust. Description of the analysis method and results of this supplementary analysis were added to the “Methods” (page 16, lines 472-477) and “Results” (page 6, lines 153-155; page 7, lines 156-158) sections in the revised main text, respectively.

However, as the reviewer noted, measurement errors in predictive factors in population-based surveys may introduce estimation bias. We elaborated in detail on the impact of measurement errors on lifestyle-related burden estimates in the “Discussion” section (page 11, lines 325-330; page 12, lines 331-341). In general, the bias may come from two sources: (1) non-differential misclassification bias caused by measurement errors of predictors in the CKB population, which may lead to underestimation of lifestyle effects; (2) overestimation of the proportion of individuals with healthy lifestyles caused by measurement errors of predictors in the CNHS population, which may underestimate the impact of lifestyle modifications on LE for the whole population. Please refer to the revised main text for detailed explanations.

Building on #7.

8) It may be helpful to add a correlation matrix of variables.

**Response:** We have added the correlation matrix plot in the revised supplementary file as Supplementary Figure 2, and interpreted it in our response to the previous comment.

9) Describe the influence of predictors on your final model. This could provide the predictors' HR (and beta coefficients). However, these can be difficult to interpret for splines and interactions; other methods could be used.

**Response:** We evaluated the relative importance of each predictor in the final model using the method proposed by Heller.<sup>6</sup> This method quantifies a predictor's relative importance by measuring changes in the model's coefficient of determination ( $R^2$ ) when each predictor is added to the model that only includes the other predictors from the final model. **Table 1** below presents the detailed results for male and female models separately, with predictors ranked in descending order of the relative importance. In both male and female models, age and systolic blood pressure had higher relative importance.

**Table 1.** Contribution of each predictor to  $R^2$ .

Men		Women	
Factor	Partial $R^2$ (%) <sup>*</sup>	Factor	Partial $R^2$ (%) <sup>*</sup>
Age at baseline	0.98875	Age at baseline	0.88243
Ln(systolic blood pressure [mmHg])	0.17023	Ln(systolic blood pressure [mmHg])	0.20239
Self-rated health	0.15934	Diabetes at baseline	0.19025

Men		Women	
Factor	Partial $R^2$ (%) <sup>*</sup>	Factor	Partial $R^2$ (%) <sup>*</sup>
Body mass index (kg/m <sup>2</sup> )	0.11318	Self-rated health	0.14115
Diabetes at baseline	0.07207	Ln(physical activity level [MET-h/day])	0.11733
Highest education	0.06361	Body mass index (kg/m <sup>2</sup> )	0.08419
Smoking	0.06352	COPD at baseline	0.05523
Marital status	0.06327	Stroke at baseline	0.04253
Ln(physical activity level [MET-h/day])	0.06255	Cancer at baseline	0.03216
Resting heart rate (beats/minute)	0.05964	Resting heart rate (beats/minute)	0.03156
COPD at baseline	0.05957	Highest education	0.03060
Stroke at baseline	0.05956	Smoking	0.02597
Alcohol intake	0.03914	Squared age at baseline	0.01893
Cancer at baseline	0.03085	Fish/seafood	0.01404
Red meat	0.00811	Fresh fruits	0.01295
Fresh fruits	0.00790	Red meat	0.00745
Fish/seafood	0.00562	Alcohol intake	0.00186
Squared age at baseline	0.00400		

MET-h/d indicates metabolic equivalent task hours per day; COPD, chronic obstructive pulmonary disease.

\*The partial  $R^2$  denotes the change in the model's  $R^2$  when this predictor is added to the model that only includes the other predictors from the final model.

10) Consider assessing calibration across subgroups - particularly groups with potential interpretation bias, such as sociodemographic groups. See recommendations in TRIPOD-AI.

**Response:** We thank the reviewer for this valuable suggestion. As the reviewer recommended, we evaluated the calibration of the prediction model across subgroups defined by education level and marital status. The calibration plots are presented in Supplementary Figures 4 and 5 in the revised supplementary file. The results showed that the model's calibration performance remained consistent across subgroups with that of the whole population. The same description has been added to the subsection of "5-year mortality prediction model developed in the CKB population" in the revised main text (page 7, lines 164-166).

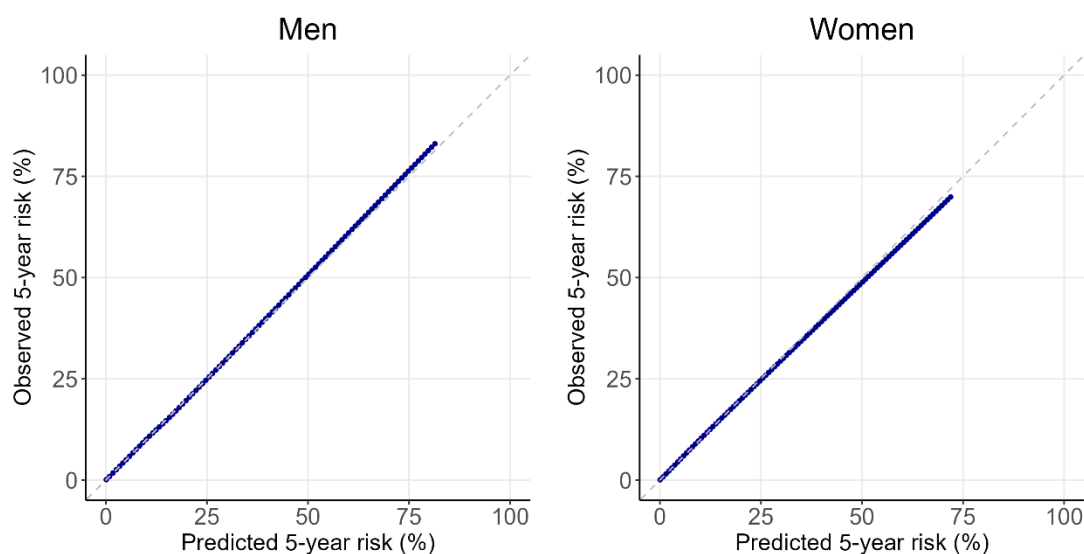
11) It is increasingly recommended that continuous calibration plots be used, but this is a minor point that does not require revision unless you are planning additional analyses.

**Response:** We thank the reviewer for this suggestion. However, we were previously unfamiliar with the continuous calibration curve. Thus, we conducted an extensive literature review and discovered a method called "HARE", which was proposed by Frank E. Harrell.<sup>7</sup> This method models the predictors and follow-up time as linear spline curves to form multiple basis functions and then combines them via maximum



likelihood estimation to derive the observed survival probabilities. We plotted the continuous calibration curves for the male and female models separately, and the curves are presented in **Figure 1** below. It shows that the observed mortality risks closely approximate the predicted risks.

We acknowledge that there may be alternative approaches for constructing continuous calibration curves, we would be grateful for more detailed guidance from the reviewer. However, as far as we know, the conventional decile calibration plots are commonly used in the calibration performance assessment of prediction models and can accurately reflect the discrepancies between observed and predicted values.



**Figure 1.** Continuous calibration plots of the 5-year all-cause mortality prediction model in the validation dataset of China Kadoorie Biobank for men and women.

12) Bootstrap validation is also increasingly recommended, but this is a minor point that does not require revision unless you are planning additional analyses.

**Response:** We thank the reviewer for this suggestion. However, in this study with a large population sample (210,203 men and 302,518 women), the simple cross-validation method is more commonly used, while bootstrap analysis would require substantial computational resources. Moreover, as far as we are aware, bootstrap analysis is more suitable for addressing the overfitting problem in small samples where splitting the dataset into training and test sets is challenging. It may also alter the distribution of the original dataset during resampling, thus introducing estimation bias.

13) Further, backward selection has noted limitations.

**Response:** As described in the “Statistical Analysis” section (page 15, lines 446-448), we used backward selection as the primary approach in the main analysis. However, we also used forward and bidirectional selection techniques to validate the results. Both

strategies selected the same set of predictors as determined by backward selection, indicating that the constructed models were stable.

14) In line 252, it seems that the equation in the math display.

The terms  $S_0(5)$  and  $\exp(V)$  should be multiplied, whereas in the paper,  $S_0(5)$  is raised to the power of  $\exp(V)$ .

**Response:** We sincerely apologize for the misunderstanding that might arise from potential inaccuracies in notation. We have revised  $h(5, \mathbf{X})$  to  $F(5, \mathbf{X})$  to denote the absolute risk of mortality in 5 years (page 15, lines 463-464).

However, to address the reviewer's concern, we also provided the detailed derivation of this formula. We calculated the 5-year absolute risk of mortality by applying the relationship between the survival function ( $S_5$ ) and the cumulative hazard function ( $H_5$ ). The specific derivation formulas are as follows:

$$S_5 = \exp(-H_5)$$

$$H_5 = H_0(5) \times \exp(V)$$

where  $H_0(5)$  denotes the cumulative baseline hazard function up to 5 years.  $V$  is the linear combination of the predictors, where the weight for each predictor represents its coefficient of association with all-cause mortality. Based on the two formulas above, we can derive that:

$$S_5 = S_0(5)^{\exp(V)}$$

$$F_5 = 1 - S_5 = 1 - S_0(5)^{\exp(V)}$$

where  $F_5$  is the absolute risk of mortality in 5 years.  $S_0(5)$  denotes the baseline survival probability at 5 years.

Reviewer #2 (Remarks on code availability):

Including the code for the model, development would be helpful, but this is still not commonly provided in papers. (I strive to publish code for our new studies.

**Response:** As suggested by the reviewer, we have uploaded the key code used in this study to GitHub, including the code for model development. The repository is available at <https://github.com/qiufen-code/lifestyle-simulation-study>.

**Reviewer #3 (Remarks to the Author):**

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

**Response:** We appreciate the time and expertise both you and your supervising reviewer have invested in evaluating our manuscript. We have carefully addressed all reviewers' comments. Please refer to our point-by-point responses attached below each comment.

**Reviewer #4 (Remarks to the Author):**

This is a good and important paper. It is for me however as a trained mathematician, not an easy paper to read. Below I state some minor recommendations which might improve the accessibility of the manuscript. I also raise some methodological queries. First I will provide an overall summary of my thoughts.

The manuscript produces novel and important results pertaining to the impact of lifestyle changes on the life-expectancy of Chinese population members. These results are of great policy interest. The authors claim that their model provides a "new analytical strategy". I do not fully agree with this and believe that their model is a nuanced application of previous work to a specific case study of China (see below). Overall I was happy with the methodological approach, although I did have a slight reservation regarding the use of Cox Proportional Hazards models (see below). Clarification either of my misunderstanding of this approach and as to why the choice of model is appropriate, or clarification of the limitations of the approach would however suffice to ease my issues. The detail of the results is a strong aspect of this work and I am more than happy with how the results are presented. Overall, after addressing my minor concerns I would be happy for this work to be published.

**Response:** We thank the reviewer for the encouraging comment. We have addressed the reviewer's concern and specific suggestions point-by-point in the following section.

Minor issues:

To begin with, after reading the paper and its emphasis on the effects that lifestyle habits have on life expectancy, I am not sure whether the authors refer to period based life expectancy or cohort based life expectancy. Are the authors assuming a stationary population with non-time-varying death rates? This should be clarified very early on as the two interpretations of life expectancy are very different, and have different uses for policy.

**Response:** We thank the reviewer for pointing this out. The life expectancy metric calculated in this study refers specifically to period life expectancy. We have made clarifications in the revised main text (page 7, line 179; page 10, line 266; page 12, line 358) and refined the methodological descriptions in the "Statistical analysis" section (page 16, lines 488-497). Furthermore, we also provided detailed explanations of both period life expectancy and cohort life expectancy, with a comparison of their uses to enhance study interpretability (page 10, lines 266-275).

I am not fully clear how this model differs/extends what has already done in the referenced model of Manuel et al (2016). You claim that your model is an analytical innovation, but I do not see how. I do believe however that your work is important as

an applied analysis specific to the Chinese population. Some other points here: you criticise the "traditional" aggregated data approach for not combining interrelating predictors, but then criticise the multivariable approach for not considering effects separately. I believe that your model's strength is in the latter approach, but I think it is unfair to be critical of Manuel et al (2016) when your paper relies heavily on that approach. Moreover, the authors should state that they apply the same methods as Manuel et al (2016), i.e., Cox Proportional Hazards to model mortality, and use a similar multivariable approach.

**Response:** We sincerely apologize for the inappropriate statement in our original manuscript. As the reviewer pointed out, this study does not constitute a methodological innovation, but rather applies this newly developed method to the Chinese population, generating research evidence specific to the Chinese population. We have accordingly revised the relevant statement in the "Discussion" section (page 11, lines 317-320). Furthermore, following the reviewer's suggestion, we have explained in the "Methods" section that this study used the same analysis strategy as the Canadian study (page 12, line 354) and removed the inappropriate criticisms of the Canadian study from the "Discussion" section (page 9, lines 251).

On page 8 you explain how Cox Proportional Hazard functions are used to calculate 5-year probabilities of death. I am not convinced that the assumption of constant hazard ratios over time is appropriate. The effects that predictor variables have on mortality risk may vary over the follow up period. It is also unclear to me whether age is a time-variant variable. Not considering (multiple) predictor variables as time-variant might mis-estimate the true effects of any one of them. For instance Body Mass Index (BMI) and blood pressure are likely to change over the follow up period and be correlated with mortality. I would like at least a discussion on this, or the authors to clarify how this concern of mine is not a problem in their model.

**Response:** We thank the reviewer for this insightful comment. As Reviewer #2 requested, we have presented the results of the proportional hazards assumption tests for the prediction model. All predictor variables satisfied the proportionality assumption, indicating that the effects of these predictors remained basically constant over time. Please refer to our response to Comment #4 from Reviewer #2.

In addition, we used baseline predictor information in the model. Nevertheless, we used follow-up time as the timescale. It has been demonstrated that adjusting for baseline age in models with follow-up time as the timescale is equivalent to adjusting for time-varying age. As for other factors in the model, such as lifestyle factors, we retained only baseline measurements primarily because repeated data collection is challenging in large-scale population surveys, and most disease burden studies do not consider time-varying variables either. However, our earlier research utilizing resurvey

data from a subset of the CKB population revealed that most participants maintained relatively consistent lifestyle patterns over long periods.<sup>8</sup> To address the reviewer's concern, we also acknowledged this limitation in the "Discussion" section (page 12, lines 332-338).

On page 10, you go through some results, and refer to Figure 2 in the supplement. I found it very hard to interpret the figure, and it did not guide me towards the significance of your findings. In particular, could you provide more detail as to what the ROC curve means/how to interpret it? Personally, I have not heard of a "c-curve", nor AUC or ROC. A simple explanation will make the paper more accessible to a wider audience.

**Response:** Due to the word limit in the main text, we have added explanations of the ROC curve and AUC in the footnote of Supplementary Figure 2 (Supplementary Figure 3 in the revised supplementary file).

In the results, I am not sure what sources of uncertainty the confidence interval cover. Are possible errors in sampling covered here?

**Response:** We are sorry we didn't make it clear. We used the same method as the Canadian study to calculate confidence intervals for life expectancy, combining the stochastic error from the model parameters and the exposure variability in the CNHS population, as described in the "Statistical analysis" section (page 16, lines 498-500). We did not consider other sources of error, like sampling error, which may lead to underestimation of uncertainty. Nevertheless, this represents one of the inherent limitations of the current method. We also acknowledged this limitation in the "Discussion" section (page 12, lines 342-344).

Grammar, use of language and other:

The first sentence need splitting into at least two. As it stands it is a poor introduction to an interesting paper.

**Response:** We thank the reviewer for this suggestion. We have segmented this statement into two distinct sentences (page 5, lines 88-92).

The second paragraph states "Healthy SG" without stating what the acronym stands for.

**Response:** We are sorry for not being clear. "Healthier SG" is a health-care reform plan launched by the Singaporean government to shift the country's health-care delivery strategy from curative care to preventive care. Promoting healthier lifestyles among citizens is a core element of this reform. We have revised the term from "Healthier SG" to "Healthier Singapore" to make it more intuitive (page 5, lines 97-98).

The manuscript could do with a re-reading and checking of all grammar.

**Response:** As suggested by the reviewer, we have conducted a meticulous review of the manuscript and corrected all identified grammatical errors.

Reviewer #4 (Remarks on code availability):

No code supplied

**Response:** We have uploaded the key code used in this study to GitHub. The repository is available at <https://github.com/qiufen-code/lifestyle-simulation-study>.

## References

1. Manuel DG, Perez R, Sanmartin C, et al. Measuring Burden of Unhealthy Behaviours Using a Multivariable Predictive Approach: Life Expectancy Lost in Canada Attributable to Smoking, Alcohol, Physical Inactivity, and Diet. *PLoS Med* 2016; **13**(8): e1002082.
2. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet (London, England)* 2024; **403**(10440): 2133-61.
3. Sun Q, Yu D, Fan J, et al. Healthy lifestyle and life expectancy at age 30 years in the Chinese population: an observational study. *Lancet Public Health* 2022; **7**(12): e994-e1004.
4. Luo Y, Li X, Li J, et al. Combined effects of smoking and peripheral arterial disease on all-cause and cardiovascular disease mortality in a Chinese male cohort. *J Vasc Surg* 2010; **51**(3): 673-8.
5. WHO-UNEP Health and Environment Linkages Initiative. Environment and health in developing countries. Geneva, Switzerland: World Health Organization; 2021.
6. Heller G. A measure of explained risk in the proportional hazards model. *Biostatistics* 2012; **13**(2): 315-25.
7. Harrell FE, Harrell F. Regression modeling strategies: Springer; 2001.
8. Han Y, Hu Y, Yu C, et al. Lifestyle, cardiometabolic disease, and multimorbidity in a prospective Chinese study. *European heart journal* 2021; **42**(34): 3374-84.



## **Response to the reviewers**

Manuscript reference number: NCOMMS-25-02875A

Title: The impact of population-wide lifestyle modifications on improving life expectancy in the Chinese population

*Note: The page number refers to that in the clean version.*

### **Reviewer #1 (Remarks to the Author):**

No further comments.

**Response:** We thank the reviewer for this positive feedback.

### **Reviewer #2 (Remarks to the Author):**

I commend the authors for their thorough and thoughtful responses to the reviewer comments. The revisions are comprehensive, and I am particularly grateful for the extensive additional analyses, including the hazard ratio correlation matrices, the assessment of potential model over-specification, and the evaluation of calibration across subgroups. These enhancements have materially strengthened the manuscript.

One comment I would add concerns the relatively low hazard ratios reported for smoking in the Chinese population. As the authors note, this is consistent with previous studies. I would suggest further acknowledging that hazard ratios are inherently relative and depend on the underlying baseline risk. In settings like China—where baseline mortality risk may be elevated due to communicable diseases, injuries, or other competing causes—the relative contribution of specific risk factors such as smoking may appear attenuated in all-cause mortality models. This does not necessarily imply a lower absolute risk from smoking but rather reflects the complexity of mortality attribution in diverse epidemiological contexts. This may also explain why smoking hazard ratios are more pronounced in disease-specific models than in all-cause mortality frameworks.

Overall, I am very satisfied with the authors' responses and revisions. I support the manuscript's publication in its revised form.

**Response:** We sincerely appreciate the reviewer's constructive comments from the previous round and are grateful for the positive evaluation of our revisions. The expert guidance has substantially improved our manuscript.

With regard to the observed relatively low hazard ratios, we agree with the

reviewer that they reflect the higher baseline mortality risk in the Chinese population. However, as evidenced by the Global Burden of Disease (GBD) data (<https://vizhub.healthdata.org/gbd-compare/>), China demonstrates lower mortality burdens from communicable diseases or injuries compared to most Western nations. This suggests alternative explanations for the baseline risk elevation. Nevertheless, participants face more other adverse factors in China, such as unhealthy environmental exposures from their homes, workplaces, and surroundings,<sup>1</sup> which collectively contribute to higher baseline mortality risk. Thus, the relative impact of lifestyle alone might be slightly diminished. The same explanation has been added to the “Discussion” section (page 10, lines 302-310) in the revised main text to further address the reviewer’s concern.

**Reviewer #3 (Remarks to the Author):**

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

**Response:** We appreciate the time and expertise both you and your supervising reviewer have invested in evaluating our manuscript. We have carefully addressed all reviewers’ comments. Please refer to our point-by-point responses attached below each comment.

**Reviewer #4 (Remarks to the Author):**

I have read with great interest the revised version of the article "The impact of population-wide lifestyle modifications on improving life expectancy in the Chinese population".

First I would like to thank the authors for investing so much energy into revising the manuscript. The article is much improved. The research is well presented and rigorously detailed. The authors have considerably improved the clarity of the study. I am now happy to recommend that the article is published.

I would like to thank the authors for addressing my previous comments:

It is important to clarify the distinction between period LE and cohort LE. The authors have now explained this. This helps explain what their results could mean for policy decision.

This research is interesting and the fact that the methods are not novel should not in any way inhibit its publication. The authors have now clarified that their model is adopted from an established framework. That being said, the results are new, and just as important. As the research heavily relies on the framework of the Manuel et al (2016) - - and to my understanding does not advance their methods in any way -- I welcome the author's removal of criticisms [of Manuel et al (2016)].

Demonstrating that the proportional hazards assumption is satisfied (log-log; discrete and Schoenfeld residual; continuous predictors) is also important, even as an Supplement. This assumption is crucial to the methods, and should not have been overlooked in the previous manuscript. I am happy with this inclusion.

I do appreciate that word restrictions constrain authors to prioritise text. However, Nature Comms is a multi-disciplinary journal, and readers (like myself) will not be versed in the terminology of epidemiology. I welcome the explanations of the ROC curve and AUC in the Supplements.

**Response:** We sincerely appreciate the reviewer's constructive comments in the previous round, which have significantly enhanced the rigor and readability of our manuscript. We are also grateful for the reviewer's positive assessment of our revisions and the final approval for publication.

**Reviewer #4 (Remarks on code availability):**

I apologise that I am not able to review the code: I am not trained in using the software STATA.

**Response:** To address the reviewer's concern, we have had our analysis codes independently verified by a STATA-qualified coauthor against the methodological descriptions in the "Methods" section. The verification results are consistent with those shown in the present study.

## **References**

1. WHO-UNEP Health and Environment Linkages Initiative. Environment and health in developing countries. Geneva, Switzerland: World Health Organization; 2021.