Response to the reviewers

Milk intake and stroke mortality in the Japan Collaborative Cohort Study - a Bayesian survival analysis

Chaochen Wang, Hiroshi Yatsuya, Yingsong Lin, Tae Sasakabe, Sayo Kawai, Shogo Kikuchi, Hiroyasu Iso, Akiko Tamakoshi

We appreciate the reviewers for their helpful comments. We have incorporated them in the revised manuscript. Please find below a detailed point-by-point responses to all comments. In the tracked changes manuscript, we provided the LaTeX diff file so that changes made can easily be identified. A clean file without tracked changes was also included.

## Comment from reviewer #1

*Thank you very much to the Editorial for entrusting me with the task of reviewing the article. The manuscript presented for review raises a very important health problem, which is stroke. Stroke continues to be one of the leading causes of death and disability in adults. It is a significant health problem which contributes significantly to reduce the quality of life. The key issue from the medical and social point of view remains effective primary and secondary prevention. The current state of knowledge indicates a significant and constantly growing importance non-pharmacological activities undertaken as part of primary prevention stroke, representing a complementary for drug therapy and just as important area of intervention. The broadly understood “primary prevention” has to be implementing it at an early stage, raising awareness of health care professionals and patients.*

*It is known that what you eat can increase or decrease your risk of stroke. The relationship between milk intake and stroke mortality is very interesting and it has a practical effect.The authors of the presented manuscript, as well as other researchers, observed that higher milk consumption is also associated with a healthier diet, greater physical activity and not smoking.*

*The summary actually shows the problem at hand.*

*The work structure is correct. Most sections are clearly and accurately described.*

*Tables are legible and correctly described.*

*The writings cited by the authors are up-to-date and appropriate.*

### **Minor revision:**

*I would make a clearer and separate section for “Conclusions” and “Study limitations”.*

*Congratulations to the authors of the interestingly described research on a large group of participants.*

[Response:] We really appreciate the reviewer for the positive comments. We have updated the manuscript with separated sections for study limitations and conclusions.

## Comments from reviewer #2

*The manuscript by C Wang et al includes a cohort of 94385 men and women from Japan, free from self-reported history of stroke, cancer, myocardial infarction and other types of heart disease at baseline who answered an FFQ regarding milk consumption. Between baseline in 1988-90 and end of follow-up in 2009, 2675 deaths from stroke were recorded on death registers. The authors applied Bayesian survival analysis to calculate hazard ratios and acceleration factors. With a range of milk consumption from “never” to “almost daily”, a higher milk consumption was associated with lower hazard and speed from dying from total stroke among men but not among women.*

*I have the following comments and questions for the authors:*

* *The introduction and discussion should however be more clear about the fact that the study is being performed in the lower range of the milk intake distribution. The dose-response meta-analysis in reference 2 indicates that the association may be non-linear. It would be good if the authors in more detail compared their results with other studies using fatal stroke as outcome and perhaps also focus on those cohorts with a low milk intake. For example, please elaborate on references 6, 20-23 in the discussion. Some of the references do not investigate milk intake per se. In the discussion, please specify that reference 19 does not investigate stroke mortality but CVD and all-cause mortality.*

[Response:] Thanks for pointing these out. We have updated the introduction as well as the discussion parts in order to better showing that these participants actually represent those in the lower range of milk intake. Discussion comparing our findings with previous studies conducted in East Asian population has been expanded in more details. The fact that reference 19 (currently reference 16) was looking into CVD and all-cause mortality has been clarified.

* *I am unfortunately not familiar with the Bayesian models used in this work. I am sure that many readers will be in a similar situation as me. I appreciate that these models can calculate hazard ratios without fulfilling the proportional hazards assumption. The authors should in more detail describe the benefits and limitations of these models in relation to the more commonly used Cox proportional hazards regression. Is a HR estimated from these models directly comparable with that from a Cox model? How is an acceleration factor interpreted? Do the models assume that a higher milk intake is associated with a lower stroke mortality, indicated also by the Pr(HR<1)? If so, how would that influence the validity of the method since you in the introduction write that there are reports of inverse associations, no association, and direct associations? In Table 2, both SD and 95% CrI are given. Is there a reason for specifying the SD?*

[Response:] Thanks for your suggestions and comments on this. Acceleration factor (AF) can be interpreted as the speed/velocity just like all individuals are driving their cars from the same start point (entry into the study setting) to their goals (stroke mortality) but different people can drive their own cars with different speeds. If the AF is smaller than 1 (reference group - never drinker), then it is considered that individual will end up at the goal with slower speed (or takes longer time). Actually these AF maybe easier to be interpreted compared with hazard ratios (HR) for readers without any training in epidemiology. Because the idea of HR includes hazard that is defined as a risk at a given point in time which is more confusing for many readers.

We understand that some readers may not be familiar with the alternative and relatively computationally expensive approach. The biggest features in the traditional Cox proportional hazard model or any other classical statistical methods are that they assume 1) there is only one true but unknown value of the parameter (in this context, the HR and AF) from the data at hand to estimate; 2) the experiment/study can be repeated, therefore the confidence intervals in that framework uses standard error to provide precision for the estimates that can be interpreted as “if we repeat the same experiment for many times, 95% of these confidence intervals will contain the true but unknown value of interest”; 3) continue on the repeatable assumption, so that the P-values are actually saying that “if we can repeat the same experiment for many times, the same or more extreme results can happen with this probability, if this probability is lower than 0.05 (commonly used as the cut-off value but without actual solid scientific support) it is considered as rarely happens or ‘statistically significant’.”

Bayesian ways of statistical analyses would disagree with all of the above assumptions because literally no one experiment/study can be repeated especially large scale cohort studies that take years and many participants/investigators to carry out. Bayesian statistics would provide a posterior distribution of the parameter that we want to estimates which means that the HR or AF is not fixed and unknown but with a distribution that has a mean and a standard deviation to describe its features, and the 95% credible intervals is directly showing us that we are 95% sure that the parameter of interest is distributed in this range. Furthermore, because of the large amount of simulations that we have conducted, the Pr(HR < 1) is showing the percentages that these HRs are smaller than 1. So a percentage of Pr(HR < 1) = 99% can be directly interpreted as “we are 99% sure that drinking milk at this frequency has lower hazard than those who never consume milk.” The models in both classical and the Bayesian frameworks are the same, so the HRs from both frameworks can be interpreted in the same way. However, the estimating processes are different: one is based on the maximum likelihood estimation but the other is based on extensive Markov Chain Monte Carlo (MCMC) simulation. Therefore we would say that the only limitation for the Bayesian approach would be the computational cost but with the development of the modern computers we expect the Bayesian approaches would become the future default. The above discussion would be out of the scope of the current paper but we understand the confusions could still exist even if we add them into the manuscript. As requested, we added some of the discussion in the revised manuscript in the “strength and limitation” section.

* *Why were analyses performed stratified by sex? Since no pooled analysis is presented, the reason for a stratified analysis should be stated in the aims, unless this was a secondary analysis and then also the results from the total cohort should be presented. The authors should also discuss the observed sex differences in more detail. Do other studies show similar sex differences? What potential biological mechanisms may explain them?*

[Response:] It is reasonable to conduct the analysis separated in men and women since we expected there is a sex difference in the association based on our prior knowledge from the same cohort as well as the Singapore Chinese Health Study. Our expectation was further confirmed through a quick likelihood ratio test comparing Cox proportional hazard models with and without the interaction terms between sex and milk intake frequencies (P for interaction < 0.0001, not shown in the main text). Our stratified analyses also suggested the AFs and HRs might be in slightly different directions among 2 out of 4 frequencies (Table 2 and Table 3, results in the “total stroke” parts), we therefore believe it is legitimate to conduct analyses stratified by sex.

The interaction effect by sex was found probably due to the fact that women in our cohort had relatively less mortality rates of fatal stroke and much better lifestyles (much less smokers and alcohol drinkers) than men regardless of milk intake frequency, or maybe due to other factors that was not available in/considered by our models such as intake of calcium or other supplements. The existing beneficial effect by milk intake is probably less evident in women than that in men in the JACC study cohort.

The updated manuscript have included these information in the methods section as well as in the discussion about sex differences on the association.

* *The aim states to “provide a more straightforward answer to the primary research question” – compared to what? And how is this more straightforward answer achieved. This is not clear, neither in the aim nor in the discussion (lines166-167). Please revise.*

[Response:] Thanks for your suggestion. The more straightforward answer is provided as Pr(HR < 1) which shows the percentages of HRs < 1 calculated through the large amount of simulations. The detailed interpretation is described above in answering your second comment. These probabilities show directly how certain it is for individuals with that milk intake frequency could have lower hazard compared with those who never consume milk. In the traditional way the answer was not clear since only estimates of HRs, 95% confidence intervals, and/or P-values were shown. None of these actually can directly answer the question “how sure it is to have lower hazard in one group compared with another?” We have added clearer descriptions about these research questions and interpretations.

* *Please describe the context of milk intake in Japan. Is milk consumed as a beverage with food, added to coffee or other consumption patterns? Are other milk products than fresh non-fermented milk consumed and could be included in the “milk” exposure?*

[Response:] Thanks for your suggestions. In the baseline year of the JACC study (between 1988 and 1990), most “milk and dairy product” consumption (92.1%) was in the form of whole milk. We have added the context of milk intake in Japan for participants who were invited into the JACC study.

* *Please specify age at baseline in abstract and methods.*

[Response:] Age range are now added in the abstract and methods accordingly.

* *Outcome ascertainment: Information on cause of death was from death certificates. Is there a possibility of individual linkage of the FFQ responses to a cause of death registry or were each of the participants’ death certificates reviewed? How many died from other causes, how many moved from the study area (and were thus lost to follow-up) and how many were censored at the end of follow-up (alive). Has stroke mortality changed during follow-up? Is there a difference in time trends of stroke subtype mortality? In how many cases was stroke subtype not recorded? The authors should also discuss potential reasons for differences in the results for stroke subtypes.*

[Response:] We have added these detailed information at the beginning of the results section. In the study area, investigators conducted systematic review of death certificates till the end of 2009. Date and cause of death were confirmed with the permission of the Director-General of the Prime Minister’s Office. Number of deaths from causes other than stroke were 18868 (10731 men and 8137 women); 5493 (2022 men and 3471 women) dropped out from the follow-up (5.8%); 67349 (25281 men and 42068 women) were censored at the end of follow-up (71.4%). Among the 2675 stroke mortality confirmed in the JACC study, 766 were unfortunately not recorded with subtypes. The median (IQR) of follow-up years for cerebral infarction mortality were 11.2 (7.1, 15.3) and 11.8 (7.9, 16.4) in men and women; the median (IQR) of follow-up years for hemorrhagic stroke mortality were 9.9 (5.4, 12.3) and 10.9 (5.9, 15.1) in men and women. Potential reasons for differences in the results for stroke subtypes are described in discussion.

* *I agree with the authors that presenting the results in the FFQ categories is a good option, rather than to calculate the amount of consumption. However, I wonder if there is some indication from the validation study as to how much is on average consumed in the higher categories of intake?*

[Response:] Thanks for your positive comments. The validation study of the FFQ used in the JACC study found that the median intake for a daily milk consumer was about 146 grams per day. This has been added in the first paragraph of “2.2 Exposure and the outcome of interest” section.

* *Why did you adjust for sleep duration, is that a factor that influences both milk intake and stroke mortality risk?*

[Response:] Thanks for pointing this out. Previous report (cited as below) using the same data showed that healthy lifestyle behaviors (including milk intake daily and proper sleep duration) were associated with CVD mortality and therefore, we also included sleep duration as one of the potential confounders.

Eri Eguchi, Hiroyasu Iso, Naohito Tanabe, Yasuhiko Wada, Hiroshi Yatsuya, Shogo Kikuchi, Yutaka Inaba, Akiko Tamakoshi, on behalf of the Japan Collaborative Cohort Study Group, Healthy lifestyle behaviours and cardiovascular mortality among Japanese men and women: the Japan collaborative cohort study, European Heart Journal, Volume 33, Issue 4, February 2012, Pages 467–477.

* *The manuscript has minor language errors throughout and would benefit from a language review.*

[Response:] Many thanks for your detailed comments that helped us to improve the quality and readability of our manuscript. We tried our best to check and fix the language errors throughout the revised version of the paper.