Reviewer's report

Title: A prospective cohort study of health behavior profiles after age 50 and mortality risk

Version: 1 Date: 3 April 2012

Reviewer: Jack Tsai

Reviewer's report:

This study examined the mortality risk associated with three health behaviors: smoking, alcohol drinking, and physical activity among older adults in 1998 and the occurrence of deaths by 2008. The authors examined the 12 different possible three-health behavior profiles that could be generated and found that smoking combined with heavy drinking resulted in a high risk of mortality especially when combined with active physical activity. The article was well-written, uses a large data set with a long follow-up period, and examines interested variables. However, I had some concerns written below.

Major Compulsory Revisions:

1. There is a lack of conceptual framework or hypotheses guiding the analyses. Examining the 12 different possible combinations of 3 health behaviors is very exploratory and these health behaviors can certainly change over 9 years (start, stop, or be episodic). What did the authors expect to find, why, and how does it contribute to the literature.

The authors suggest their approach has not been used before, but the relationships between the health behaviors examined have certainly been examined in similar ways in other forms (e.g., studies have shown the harmful additive effects of smoking and drinking, smoking and drug abuse, obesity and smoking, etc.).

2. Do the results show a higher risk of mortality with a smoking, physically active, heavy drinking profile than the same profile except they are inactive? Am I understanding that correctly and why do you think that is? The abstract does not reflect that finding, but instead suggests smoking “overshadows” inactivity.

3. Authors did not examine time of death, that is, how much time after 1998 did they die? Also they did not control for various important sociodemographics like income, which may determine quality of health care and odds of survival from health problems.

4. In the Methods, the authors say their sample is representative of adults older than age 50, but merging several samples does not necessarily result in a representative sample without weighting procedures or some other method to overcome non-response bias and other sampling biases. The authors mention “HRS population weights” although it is unclear what they are (e.g., relative weights?) and whether the sample remains representative if participants who had incomplete data were excluded.
5. In the Methods, there is a lack of transparency about how the final sample was obtained. How and what additional cohorts were sampled? What was the original number of people invited to participate, and what was the original final sample of participants before excluding those who had incomplete data (i.e., how many were excluded because of incomplete data)?

6. In the Methods, the Health and Retirement Study should be described in more detail, and a citation should be provided. Was it a telephone survey? Mail survey?

7. In the Results and Discussion, given the size of the sample, statistical significance should not be focused on because even minute differences will be significant. In fact, I suggest all statistical tests be conducted at least at the p<.01 level if not higher. The magnitude of the odds ratios should be a focus instead of statistical significance.

8. In the Results regarding Table 1, authors make “eyeball” comparisons between the 12 different profile without accounting for variability in the measures. If the authors do make profile comparisons with statistical tests, they need to somehow control for the large number of comparisons made.

9. Tables 2-4 are quite difficult to read and interpret, and are not reader-friendly. For example, in Table 2, I am not sure what the two rows are for each label, e.g., current smoker has a 2.0 and 1.9 for the full sample? Both tables need to be redesigned to be easier to read and the authors should focus on the main findings to display.

Discretionary Revisions:
1. Why was the sample stratified and why into these two age groups (51-65 and 66+)?

2. I’m not sure I understand why there initial models that adjusted for sociodemographics and then fully adjusted models controlling for sociodemographics and baseline health problems? Why not only present the fully adjusted models?

3. There should be some discussion in the Introduction about why these three health behaviors were selected and literature should be provided discussing how these three behaviors are often correlated, i.e., those who drink heavily are likely to smoke often and be inactive.

Minor Essential Revisions:
1. Not sure self-rated health needs to be abbreviated to SRH as it hinders readability slightly

2. The right part of Figure 1 is completely missing, which is a large oversight.

3. Not sure some of the row labels need to be italicized for Table 1. Some numbers in Table 1 are italicized, and it’s not clear if it was a formatting error or for denotation.

Level of interest: An article whose findings are important to those with closely
related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.