Reviewer's report

Title: A retrospective, cross-sectional study of real-world values of cardiovascular risk factors using a healthcare database in Japan

Version: 1 Date: 21 April 2014

Reviewer: Yuling Hong

Reviewer's report:

This manuscript has addressed a very important question of using real world data from employment-based health insurance companies for possible future surveillance and research purpose. The data were massive and the study was in general well executed. However, this reviewer has a couple of general comments and several specific ones for the authors to consider.

General comments:

1. In addition to smoking and biological factors related to CVD, were data on other behavior factors such as physical activity and dietary intake collected by MinaCare as well?
2. You mention that CVD risk factors are not controlled several times, but I do not see that you reported any data on treatment and control. You had data on prevalence only. Suggest that to use a different word than “controlled”.

Specific comments:

1. Page 3, abstract: the Background section is longer than all other sections. Suggest shorten than section and beef up the Results section.
2. Page 3, abstract, the Results section: what was the rationale to highlight only results from those aged 55-59 highlighted here?
3. Page 8, lines 9-12: the reason (with most of the subjects) of choosing 2011 data was not strong. Possibly be better to say that it was chosen to make it comparable with two other National Survey Data as you used 2011 NHLW-H&N data in the study. On the other hand, you included 2010 MHLW-SH data. This makes me think why you did not use both years of 2010 and 2011 data from the MinaCare database.
4. Page 8, lines 15-16: it would be important to describe the frequency of the missing values for key study factors in the text or a table so the readers can see the completeness of these measures in the database. If the proportion of missing values is high, you need to discuss this as a limitation of the database.
5. Page 9: it would be good to mention which guideline was the cutoff 140 for LDL-c based on under the second paragraph.
6. Page 10, 1st paragraph: it may be better to be moved up after you described the MinaCare database.
7. Pages 11-17: it is difficult to follow when you describe the differences across
age groups and also to understand why data for certain age groups were reports. You may want to re-organize the presentation of the results for clarity.

8. Page 15 under the BP subsection: it would be good to mention whether the same BP measurement protocol was adopted across the three databases/surveys. For example, the number of BP measures was used. You might have only one measurement for MinaCare but more for MHLW-H&N, for example.

9. Page 18, line 1: I would suggest deleting the word of “full”.

10. Page 19 before the Conclusion section: it would be good to mention the limitations of the MinaCare database.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

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

Declaration of competing interests:

I declare that I have no competing interests.