Reviewer's report

Title: Tamoxifen Use reduces Osteoporotic Fracture Risk in Asia Breast cancer woman: A nationwide population-based cohort study

Version: 3 Date: 22 March 2015

Reviewer: Lisa Lix

Reviewer's report:

This is an interesting manuscript conducted in a large population-based dataset. However, there are multiple important elements of the methodology missing from this manuscript.

Major Compulsory Revisions:

1. The authors report that they recruited subjects between 2000 and 2011. It is unclear how the duration of follow-up times might have differed for the Tamoxifen and non-Tamoxifen user groups. The authors must provide information on the mean and median follow-up times. Also, given that the recruitment period extended to 2011, and the follow-up time ended in 2011, it seems possible that some individuals had no observation time for fracture. Please provide additional detail to address this point.

2. It is not clear how the exposure (i.e., Tamoxifen use) was measured as a time-dependent covariate, and at what intervals these measurements were taken. Also, the authors should provide the list of the Anatomical Therapeutic Chemical (ATC) codes used to indicate Tamoxifen use. Please provide clear information in the Methods section about the unit of measurement for Tamoxifen use.

3. Using administrative health data to identify individuals with breast cancer requires the use of a validated algorithm for case ascertainment. Please provide information about the sensitivity and specificity of the algorithm. High specificity is required to decrease the chances of misclassification of cases. As well, six comorbidities were included in this analysis as confounding covariates. Please provide a justification for their inclusion and address the potential for misclassification bias in the diagnoses used to ascertain these comorbidities. It is unlikely that the diagnoses will have equivalent sensitivity and specificity for case ascertainment. Finally, the manuscript lacks important details about the sensitivity and specificity of the diagnosis codes for osteoporosis-related fracture. As well, if they authors are interested in capturing major osteoporosis-related fractures, they should have included diagnosis codes for fractures of the humerus.

4. Why did the authors not use a propensity-score model to adjust for confounding? Also, potentially important variables such as cancer stage, socioeconomic status, and rural/urban residence, which could all affect fracture
occurrence and Tamoxifen exposure, are absent from the model. There is the potential for confounding bias in the hazard ratio estimates that were obtained.

5. The authors make the following statement in the Discussion section: “Our study doesn’t compare the effect of other aromatase inhibitors with Tamoxifen. The current results revealed the protection effect was not related to the duration or dosage. It may then be due to the steady usage of Tamoxifen once a patient has prescribed this agent in Taiwan.” This statement is difficult to interpret without the authors providing clear explanations of how dosage and duration of Tamoxifen use were measured.

6. The Results section is missing important information about model fit.

Minor Essential Revisions:
1. The authors mentioned testing for gender differences in the Methods section (page 6). This seems to be an error, as the study was supposed to be conducted amongst women only.

2. Table 1: Please include column headings to distinguish frequencies and percentages. Also, provide a definition of the acronym PAD in the table note. Using both the chi-square test and Student's t-test for age is redundant. It is preferable to treat age as either a continuous or categorical variable, but not both types of variables. Finally, given that the p-values were all less than 0.0001, it would be more efficient to report in the table note that all differences between exposure groups were statistically significant at alpha = .05.

3. Table 2: Please provide a note to define the acronym DDD. Also, include “HR” in bracket after “Hazard risk” in the table title as is done for the 95% confidence intervals. There is redundancy in reporting both 95% CIs and p-values in the same table to indicate statistical significance. The latter should be excluded.

4. Table 3: Please provide a note to define the acronym PAD. Similarly, include “HR” in brackets after “Hazard risk” in the table title as is done for the 95% confidence intervals. Again, there is redundancy in reporting statistical significance with 95% CIs and p-values.

5. Table 4: Please include column headings in the table to indicate hazard risk and 95% CIs. Again, there is redundancy in the reporting of statistical significance with 95% CIs and p-values.

6. Table 5: Same comments as for Table 4.

Discretionary Revisions

None

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests.