Reviewer’s report

Title: Use of comorbidity lookback to determine prevalence and risk in pregnancy

Version: 1 Date: 3 February 2011

Reviewer: Hangsheng Liu

Reviewer’s report:

Summary

The authors used comorbidities identified in multiple years of hospital data to predict obstetric hemorrhage among 53,438 women who gave their first birth during 2005-2006 in New South Wales, Australia. They examined (1) whether a longer lookback would identify more chronic conditions using hospital data; and (2) whether a longer lookback would improve the model performance in predicting obstetric hemorrhage. The authors found that most chronic conditions could be identified using the hospital records in the prior 2 or 3 years. But the additional comorbidities from a longer lookback did not improve the model performance in predicting obstetric hemorrhage.

Comments

The authors raised an interesting question and the research community will benefit from the results. There are some concerns that need to be addressed to ensure the robustness of the results.

Major Compulsory Revisions

1. It is not clear to me based on the manuscript, why this healthy population and the outcome (obstetric hemorrhage) were selected to answer such a question. In other words, this population and the outcome (hemorrhage) may not be sensitive enough to the inclusion of more chronic conditions, which may explain the results. As shown in Table 1, the number of individuals with an admission in prior 2~5 years is really small compared to the overall sample. Therefore, not much additional information could be gained using a longer lookback. At minimum, this should be listed as a limitation of the study and should caution the readers on this.

2. I think it is critical to have a complete count of chronic conditions. The underestimation of the prevalence of some of the conditions may be due to lack of outpatient data? Maybe I missed this when reading the manuscript. It would certainly be a great improvement if outpatient data were included. A relevant point is that the women living in NSW may be admitted to hospitals outside NSW. If no outpatient data, it should be mentioned as a study limitation.

3. It is unclear to me what the sample size is for each model in Table 4, which is the most important table. If all the models have the same sample size (53,438), it
is likely that the reference group varies across different models as more individuals are identified to have chronic conditions as the length of lookback increases. I wonder if it would be worth using the same reference group for all the models, i.e., the individuals without any chronic conditions in all the years. Also, the authors may want to have a statistical test to compare the C-statistic across different models.

4. The authors spend a lot of space on discussion etc., but the key results are missing. There are only 2 models in Table 4. I am not sure why other models for 2, 3, 4, and 5 years are not presented?

5. The authors used stepwise logistic regression and 2 conditions were finally entered into the final models. I am not sure if the stepwise selection would gain much as the sample size is huge and there are only 8 disease categories available. One possibility is that the authors may force all these 8 diseases into the model while using the stepwise selection method for other covariates.

Minor Essential Revisions

1. The authors may want to further justify why these chronic conditions were selected. The solution might be to cite prior literature on predicting obstetric hemorrhage.

2. The first paragraph in the results section is very confusing. One way is to present the results not shown in Table 1 or Figure 1 in a separate paragraph. The sentence starting with “In addition to the birth admissions, 1,831…” is not clear (maybe because a typo near “both the birth and record subsequent”?). The number “55,269” does not appear in either Table 1 or Figure 1. After many reads, I figured it was the summation of 1831 and 53438. Also, Figure 1 presents some exclusions but nothing is mentioned in the text.

3. The second paragraph, the last sentence, “6.8%” should be 6.4% (=1-93.6%) according to Table 2.

4. Not sure why Table 3 only presents 5 of 8 conditions. Is there any reason? Maybe I missed something.

5. It may be implied but is not explicitly stated in the text or tables, whether a longer lookback includes all conditions found in a shorter lookback. For example, in Table 3, does “Pregnancy” column include all conditions found in “Birth” column and so forth? This is directly related to the conclusions.

6. The authors do not explain why they didn’t use comorbidity index and why they didn’t construct two chronic condition variables, one for the index admission, and the other for prior admissions.

Discretionary Revisions

1. The title could be more specific. It is not clear now what “prevalence” or “what risk” the authors are trying to estimate or evaluate.
2. In the abstract, Methods, the authors may want to spell “NSW.”

3. Table 1. It may be better to present “at birth admission” first, then “within the birth transfer sequence?” I am not sure about the exact definition of “within the birth transfer sequence” though.

4. Figure 1, just showing the number of patients would probably make the figure easier to follow.

**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.