Reviewer's report

Title: Can we use the pharmacy data to estimate the prevalence of chronic conditions? An attempt to compare different information sources

Version: 3 Date: 6 June 2011

Reviewer: Lisa Lix

Reviewer's report:

Overall Comments
This study provides an interesting comparison of different population-based data sources to estimate the prevalence of chronic conditions for a single geographic area. The key objective is to assess whether a method for ascertaining cases based on pharmacy data produces estimates that agree with the estimates from other administrative and self-report data sources.

While I believe that the study has merit, the manuscript requires substantial revisions in order to ensure that the proposed methodology of ascertaining cases from pharmacy data is valid and generalizable. I was not involved in an earlier review of this manuscript. While I understand that revisions have been made based on previous reviewers’ comments, there are still a number of changes that I would recommend.

Major Compulsory Revisions
As noted by a previous reviewer, categories that are not define specific chronic conditions, including pain, transplantation, and hyperlipidaemia, should be removed from Tables 1 to 3.

Please remove from Tables 1 to 3 all conditions for which pharmacy data are the only source of a prevalence estimate, including gastric acid disorders, benign prostatic hyperplasia, anaemia, gout, cystic fibrosis, growth hormone deficiency, and erectile dysfunction. It is not meaningful to report chronic conditions for which there is no comparator data source.

Stratify the analyses by sex, so that separate estimates are reported for males and females. Given that some of the conditions (e.g., osteoporosis) have very different prevalence estimates for males and females, this stratification provides one tool to assess the validity of the pharmacy data methodology.

The authors must provide a description of the survey methodology and the questions used to ascertain chronic conditions from survey data. This information should be included in the section entitled “Data sources”.

Confidence intervals must be provided for all prevalence estimates. The method used to produce these confidence intervals should be described in the Methods in a new section entitled “Statistical analysis”. The authors should also test for
differences in the estimates produced using each data source and describe the
method used to conduct these tests in the section entitled "Statistical analysis".

Minor Essential Revisions
The authors indicate that when estimating prevalence from pharmacy data they
excluded individuals with short-term use having less than “three boxes” during
the year 2006. Please define the term “box”. Was a sensitivity analysis
performed, to determine if increasing the number of boxes resulted in substantial
changes in the prevalence estimates? If the authors did not conduct a sensitivity
analysis, this needs to be described as a limitation of the study.

Discretionary Revisions
The interpretation of the study results would be easier if the authors graphed the
differences in estimates for pharmacy data and the other data sources.
Confidence intervals for the differences should also be provided.

Change the title to “Can we use pharmacy data to estimate the prevalence of
chronic conditions? A comparison of multiple data sources”

**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.