Author's response to reviews

Title: Can we use the pharmacy data to estimate the prevalence of chronic conditions? A comparison of multiple data sources

Authors:

Francesco Chini (chini@asplazio.it)
Patzio Pezzotti (pezzotti@asplazio.it)
Letizia Orzella (orzella@asplazio.it)
Piero Borgia (borgia@asplazio.it)
Gabriella Guasticchi (guasticchi@asplazio.it)

Version: 4 Date: 13 July 2011

Author's response to reviews: see over
Natasha Mellins-Cohen  
Head of Publishing Operations  
BioMed Central  

Rome, July 13, 2011

Dear Editor,

We would like to submit the revised version of the manuscript MS 6066988615209581. Please note that the new version follows the copyediting provided by the first reviewer and includes a figure with results stratified by sex, as suggested by another reviewer. The results now refer to 20 instead of 31 chronic conditions excluding all categories for which pharmacy data are the only source of a prevalence estimate or the category was not a specific chronic condition. Also the title is slightly modified to take into account the suggestion of one reviewer. Furthermore, during the revision process we found that for the registry of exempts some cases were erroneously excluded from the analyses. It is of note that the results are still similar and this does not change the interpretation of the data. Finally, attached is a point-by-point list of answers to the reviewer’s comments and in the new version the changes are coloured in red.

Please note that:

• All authors declare that there are no competing interests.
• All of the authors listed concur with the content of this paper and its submission to your journal. The final version of the manuscript has been read and approved by all authors.
• Neither the article, nor any part of it, is under consideration for publication in any journal.
• The study design did not require an approval by an Ethical Committee. The Agency for Public health is the mandatory collector of all the data used in this paper.

On behalf of the authors

Francesco Chini
Reviewer: Marc Saez

Reviewer’s report:
Major Compulsory Revisions

It is not clear to me how authors 'aggregate' 'general population' (prescription drug usage) and a 'selection sample' (hospital discharge registry). Authors should adjust these two samples: the former, at least theoretically random, the latter, with a clear selection bias, before proceed to estimate prevalences and do not, simply, discuss this point as limitation. In this sense, authors should read the paper Saez et al. A selection-bias free method to estimate the prevalence of hypertension from an administrative primary health care database in the Girona Health Region, Spain. Comput Methods Programs Biomed. 2009;93(3):228-240. Authors should apply and/or the methods used there. Although we theoretically agree with the comment of the reviewer, we remark that the main goal of our study is to provide a crude comparison of prevalence estimates for several chronic diseases using “without adjustment for selection bias” different routine health databases available in Italy and thus in the Lazio region. As we specified in our discussion section, for a specific chronic disease, methods that combine different health databases (e.g., capture-recapture techniques) or statistical methods that adjust for selection bias using external information obtained by surveys can provide more reliable estimates. However, these techniques can be applied only if you can link health databases or if you have external data. We did not have the external data to apply the technique suggested by the reviewer and we were not authorized to link the health databases due to privacy reasons.

Furthermore, these simple estimates are also compared to those obtained by the survey conducted by the National bureau of census (i.e., ISTAT) and we evidenced that in some cases the survey estimated a lower prevalence because of sampling error or other problems to identify people with that condition.

We revised the limits section to include the paper of the reviewer and the possible techniques to adjust for selection bias.

It is not clear how authors try to reconcile actual prevalence rates and prevalence rates based on diagnosed people (maybe coverage of drug treatment is low for some health problems, not all the problems are equally controlled, etc.). It is not enough the discussion. Authors should extend more this point.

As we wrote in the previous point the objective of the study is to provide a comparison of prevalences of people diagnosed with several chronic conditions. However, we added as first limitation the fact that the prevalence rates of diagnosed people can many only partly reflect the actual prevalence rate. We highlight that some of the points suggested by the reviewer were already reported in the limits section.

Some drugs could be used for more than one health problem. How authors controlled for this limitation?
We added in the methods section a sentence that better explains how some drugs refer to a specific chronic condition. More specifically the following paragraph reports how we controlled for drugs used for more than one health problem:

“Drugs for CCs treatments might be totally or partially reimbursed by the RHS and are often subject to restrictive note (in Italian called “Nota CUF”) for dispensing defined by the Italian Medicine Agency (AIFA – Agenzia Italiana Farmaco) (18). These restrictions can be considered as guidelines for a more appropriate use of pharmaceuticals. The “Nota CUF” defines the CC for dispensing the drug and increased our ability to capture drug users affected by the selected CCs.”

The restriction to “Nota CUF” is reported also in table 1 in the footnote.

Discretionary Revisions
Maybe authors should add some figures (for instance, instead/in addition of Tables 2 and 3).

In agreement with another reviewer (Lisa Lix) we add a figure stratifying by sex the previous estimates.

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.

We agree with the reviewer comment and we removed the specific indicated categories.

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.
Following the suggestion of the reviewer we also removed these other chronic conditions for which there is no comparator data source. After these removals we have 20 categories.

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.

We stratified the estimates also by sex, as suggested. These are reported in figure 1(A and B). We added a sentence in the method section and accordingly we revised the results section.

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”.

Ok. We added the section entitled “ISTAT Health survey 2004-2005”

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”.

For the survey we provided 95% confidence intervals (95%CI) for the prevalence estimates. We also added a sentence in the methods section to describe the way we calculated the 95%CI in the prevalence estimates section. Regarding the prevalence estimates obtained by the health administrative databases we cannot provide any 95%CI because they are based on the entire population of Lazio region. We do not feel that a new section entitled “Statistical analysis” is necessary. Finally, given that estimates obtained from the different health databases are not based in probabilistic sample there is not any reason to provide any formal inferential test to compare differences.

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.

We changed the term “box” with “package”. It is not clear to us how to better explain the meaning.

A sensitivity analysis was not performed. We added a sentence in the limit section.
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.
We added as previously suggested by the reviewer, some graph stratified by sex. Graph with differences and confidence interval are not provided for the reason explain At the previously point.

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

Reviewer: Ronald Cossman
Reviewer's report:
Major Compulsory Revisions

How do you reconcile the differences between prescription drug usage, which is applied to the entire population to determine a rate, versus the hospital discharge and registry of exempts, which is limited to those seeking attention by a medical professional and/or medical facility? In other words, the denominator is not the same. Either adjust such that you can compare to Rx rates or drop the comparison entirely.
Our primary goal was to provide a simple comparisons of the diagnosed prevalence of several CCs estimated in a crude way using different health administrative databases. We expected that for some CCs the hospital discharge registry and the registry of exempts strongly underestimate the actual diagnosed prevalence of the CC and this was discussed in the discussion section. We revised the title of the manuscript trying to take into account this comment and comments of the other reviewers. Following a suggestion of another reviewer we added a reference about a method to correct the prevalence for selection bias when using hospital discharge databases.

One methodological issue that is not addressed is the difference between actual prevalence rates and the rate of those who have been diagnosed and are being treated with a drug regimen. See: http://www.pophealthmetrics.com/content/7/1/16 as a starting point.

Following the suggestion of the reviewer we further revised and enlarged the limitations section highlighting the difference between actual prevalence rates and prevalence estimates evidencing in particular those CCs for which the coverage of drug treatment is low. As previously described we also added another reference suggested by another reviewer.
Minor Essential Revisions
The manuscript could be improved and polished by a line edit for word choice and
construction.
We received by the editorial office a copyedited revision of the manuscript and we revised
the current version starting from that revision. We have now submitted a new version with
tracked changes in red.

Discretionary Revisions
I would prefer to see the equations/process by which the Rx calculations were made.
We feel that the prevalence estimates do not require any description of the
equation/process. In case the editor will request it, we can provide formal calculations.

I would like to see a discussion of the applicability of these rates to the general public
and/or limitations to this methodology.
The discussion was entirely revised following the reviewer’s suggestions and particularly
the limitations section

Are there other sources of Rx other than the pharmacy?
In Italy and specifically in our region there are a couple of health administrative databases
that we can mention: the emergency departments visits registry and the outpatient
specialty care registry, which includes all diagnostic and therapeutic procedures performed
in Lazio region. However, both were not considered in this study because the information
collected in these registries did not permit to identify with good accuracy people with many
of the CCs here evaluated. We feel that it is not needed to add this issue in our
manuscript.

How do these rates compare to published rates (regardless of data source) in other
regions of Italy?
The prevalence estimates we provided are consistent with others previously reported in
Italy and in US and we added in the second paragraph a sentence about this point.