Reviewer’s report

Title: Disease identification based on ambulatory drugs dispensation in comparison to ICD-10 diagnoses

Version: 2 Date: 17 June 2013

Reviewer: Björn Wettermark

Reviewer’s report:

Thank you for the opportunity to read this interesting paper. I was not involved as a referee from the beginning and had the task to respond to the previous comments. It is a challenge since there are areas where we are not in complete agreement, and since there may be previous questions raised that I don’t completely capture. In general I found the paper of interest and I acknowledge that you have taken referee comments into account in the revised version. Still I have some questions around the context and the methods, their applications and generalisability as discussed below.

Major Compulsory Revisions

Aims, page 5: It is claimed that “We chose three distinct fields of application: morbidity indicators, ambulatory care cost control, and health insurers’ risk adjustment.” This may be relevant, but the usefulness for these purposes has not been studied. Consequently, I think it is good to discuss around it but that it shouldn’t be an aim. This relates also to the conclusion, page 17: The usefulness of this data for health insurers’ risk adjustment or ambulatory cost control has not been studied and, consequently, it should not be mentioned as a conclusion. The conclusion in the abstract, page 2, is not either synchronized with the general conclusion and supported by results.

Background, second paragraph on page 3: “Outpatient morbidity information is however scarce, except for cancer registers and contagious infections...” is not true. There are many scientific papers published on prevalence of common diseases such as asthma, COPD, hypertension, heart failure, atrial fibrillation... These studies are based on data from registries, electronic medical records and/or population surveys. I would suggest the authors to conduct a more thorough literature search on disease prevalence

Background: The statement “Most other countries have abandoned such records, mainly because data collection is time-consuming, costly and not always reliable” refers to a 15 years old paper. The electronization of healthcare has changed the situation completely since that paper was written with diagnoses electronically recorded in ambulatory care in many countries. This is presented in many recent reviews such as:

Methods. Page 5: I am uncertain what you mean with the sentence “As in most other developed countries, only diagnoses which had an impact on the treatment of the patient were collected [21].” This reference just refers to the ICD-instructions? If this really is the case, then it should be supported by references.

Methods, page 6: There is still some uncertainty on how drugs and diagnoses are linked together in time. It is claimed that “to avoid data correlation, patients were considered only once.” But what diagnosis was taken, the one most closely in time?

Methods/discussion: The studied population included 108,915 insured enrollees followed throughout 2005 and 2006; they were hospitalized at least once. It is important to keep in mind that the population was highly selected. This is partly addressed in the discussion, but I would like to see also a section in the methods about the context, i.e. the representativity of those included in the scheme and the study and perhaps also some general comments on the Swiss regulatory and reimbursement systems for hospital consultation and drug use. This is important to better interpret the study findings to other countries.

Results, page 12-13: Some results are presented by context of use: morbidity indicator (M), ambulatory costs control (C), insurers’ risk adjustment (R) only for chronic conditions further groups for ambulatory cost adjustment (letter C). It is also reported that twenty three were not retained because of their poor accuracy or their treatment was prone to practice variations.. This analysis and how these categories were assessed ought to be mentioned in the methods section.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Title: I suggest “inpatient” to be added to the title to make it more clear that the study is limited to this population. This could, e.g., be: “…in comparison to ICD-10 Diagnoses in inpatient care”

Key words: There are several more appropriate key-words than only “case mix”, “pharmacy data”, “ambulatory care” The terms chosen should more reflect the actual study.

Results page 13: Table 3 shows the reproducibility of drug information based morbidity categories from one year to the next. This should refer to table 4 instead. Later on the Morbidity prevalence rates (crude and adjusted) inferred from 2005 drug information for the prevalence rates should refer to Table 5 instead of Table 4.
Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

Abstract: The first sentences “The relative unavailability of outpatient diagnoses data makes pharmacy-based case mix measures attractive. Most published tools use national drug nomenclatures and lack head-to-head comparisons of drugs-related categories with their diagnoses-based analogues.” are difficult to understand. They could be made more clear for the readers.

Methods. Face validity of each morbidity category inferred from drug information was ensured by thorough clinician and pharmacist reviews. It would be nice to describe how this was done more in detail.

Methods, page 7: The sentence “Yet we know that co-morbidities are often not recorded in hospital minimal data sets…” more relates to the discussion than the methods

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