Reviewer’s report

Title: Methods for identifying 27 chronic conditions: Application to administrative data

Version: 3
Date: 29 December 2014

Reviewer: Corinne Hohl

Reviewer’s report:

Major Compulsory Revisions

Major Revisions:

1. What was the rationale for selecting the 40 health conditions you sought algorithms for? The rationale for their selection should be justified clinically. If this list was merely taken over from previous authors, the previous rationale for selecting these should be presented.

2. Some of the conditions that met your criteria can hardly be regarded as significant from the perspective of contributing significantly to morbidity, mortality and cost (e.g., irritable bowel syndrome) in comparison to others (e.g., metastatic cancer), while other additional conditions might be very important to consider (e.g., HIV). Please explain.

3. Methods, line 101 and beyond: The search strategy and databases you searched fall short of a systematic strategy. It is likely that relevant studies with algorithms that might have been of interest to your group were missed.

4. No qualitative assessments of any studies was reported. This is a shortcoming, and should be reported. In particular, it is relevant how different coding algorithms were validated by primary studies. I am unable to read the header row in Table 1 – so this information might be there. Ideally, this information would be provided qualitatively in brief phrase in a tabular format.

5. Methods, lines 111-112: Please provide a reference.

6. Methods, lines 139-141: I am unclear of the purpose of demonstrating “feasibility” which is quite different than demonstrating “proof of concept”. Proof of concept implies verifying that a concept or theory has the potential of being useful for a specific purpose. Simply applying an amalgamation of coding algorithms to administrative data is in my view insufficient to demonstrate that this yielded meaningful data that can be used for research or clinical purposes. There should be parameters, defined a priori, based on clinical knowledge of disease prevalence, that would give you the ability to test whether or not the application of the algorithms yielded meaningful results in your cohort.

7. Results, lines 175-176: “Second, we combined the ‘highly likely’ and the ‘likely’ codes from the original algorithm for chronic pain”. Please explain. Why was this
done, and for which algorithms? Many of the conditions you mention are not expected to cause pain.

8. Results, lines 169-187: This should be in the Methods section if your intent was to develop, refine and test new coding algorithms identified from your literature search. If so, please clarify your objectives, and place this paragraph in the Methods.

9. Results, lines 192-193: The statement “Approximately half of all participants did not have any of the 30 morbidities for which high or moderate validity algorithms existed” can simply not be supported. It might be correct to state that your algorithms were unable to identify conditions in these patients, but just because none of your algorithms triggered a positive result does not mean that these patients did not in actual fact have any of these conditions (see your data on kidney dysfunction). The coding might be inaccurate, your algorithms insensitive, etc. I can only support this statement if you can present another standard for determining whether or not these patients had these diseases. Please rephrase this statement, and following statements to accurately reflect what you showed.

10. Results, lines 204-208: This is a great example of validating a coding algorithm. In this case however, finding such a large discrepancy (five fold!) in the actual versus coded prevalence of kidney disease should lead you to conclude that the administrative data and coding algorithms with regard to kidney dysfunction are insensitive and not likely to be useful. It is likely that their use would yield a gross underestimate of their prevalence. I’d use these results to argue that you should not be using your suggested approach for identifying multimorbidity. This finding should make you call into question the validation approach used for this algorithm – and hardly supports the finding of “adequate sensitivity” (Discussion, lines 246-247).

11. Discussion, Lines 221-223: I disagree with this conclusion. Present ways to validate your prevalence estimates in this cohort, and if the prevalence of multimorbidity falls within an a priori defined equivalence range of what the multimorbidity prevalence should be based on an independent standard (registry data, chart review of a small selection of randomly selected patients, etc.), this statement could be supported. Validation of any algorithm, and also of refinements of algorithms, should be undertaken prior to their use.

12. Discussion, lines 235-237: Please present a table in which you tell the reader how these codes were validated.

13. Discussion, line 260 and beyond: I strongly urge you to take the time to validate your algorithms for Canadian administrative data before you continue to use them.

14. Discussion, lines 270-271: I fully agree – use of medication dispensing data may have allowed you to validated coding algorithms for conditions which mandate treatment (e.g., insulin dependent diabetes).
15. Discussion, line 281: Extensive use of sensitivity analyses is rarely if ever a good idea. I would remove this statement.

16. Table 1. The header row is not legible. Please describe methods used for validation of the individual algorithms.

17. Table 1. You state in the table header that all algorithms are validated, but you do not report diagnostic test characteristics (PPV, sensitivity) for all. A coding validation study should report the diagnostic test characteristics of the codes studied, and I thought all selected algorithms you chose had been validated. Please clarify.

18. Table 2, second row: Please remove the row with “none” morbidities, it is counterintuitive here in relation to the categories listed below.

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.