Reviewer’s report

Title: Correlating pharmaceutical data with a national health survey as a proxy for estimating rural population health

Version: 1 Date: 3 November 2009

Reviewer: Greg Roth

Reviewer’s report:

Major Compulsory Revisions:

1. I am concerned that your reliance on IMS’ NDTI data leads to a high degree of misclassification of both medications and disease categories. You measure the number of "heart disease" prescriptions filled in a year in a state and compare this with individuals who report being told that they have "coronary heart disease". Your approach is complicated by the fact that, while some of the prescription medications you are counting are used for the treatment of individuals with coronary heart disease, all of them are used for the treatment of other, more prevalent conditions. Diabetes is relatively straightforward because diabetic medications are used, for the most part, only by diabetics. For stroke and heart disease, however this is not the case. For instance, the most common medications used for the secondary prevention of events in individuals with CHD (ASA, beta blockers, ACE inhibitors, and statins) are also the most common medications used for primary prevention of CHD in patients who do not have CHD. Individuals receiving them would not have been told by their health provider that they had CHD and likely would have been reassured, specifically, that they should take these medications to avoid developing CHD. This includes the vast number of individuals who are hypertensive (and therefore placed on beta-blockers, ace inhibitors, angiotensin receptor blockers, and calcium channel blockers) or hyperlipidemic (and therefore placed on lipid-lowering agents such as statins). Furthermore you count alpha blockers, which are most commonly prescribed at low dose for benign prostatic hypertrophy and only rarely used for hypertension anymore. This leads to a further concern, namely that you may not have, as you state on page 9, "paired drug classes with appropriate BRFSS questions". For instance, you estimate correlations between heart disease prescriptions (which include a large number of lipid-lowering medications) and individuals with high blood pressure (a condition not treated with lipid-lowering).

2. NDTI offers a confusingly broad list of medications for cerebrovascular disease, including: oral anticoagulants (i.e. warfarin) that are actually used for atrial fibrillation and only rarely to treat individuals who have suffered a stroke, oral-antiplatelets (aspirin, clopidogrel and dipyridamole) that are used widely for a many different conditions including coronary heart disease as well as stroke, Vitamin K which is actually used to reverse anticoagulation, as well as dementia medications, anti-Parkinsonians, and anti-epileptics. If your methods work, this list will still lead you to dramatically overestimate the prevalence of those having
suffered a cerebrovascular accident (stroke). Furthermore, it is unclear why you would compare this population with individuals who have hypertension (Table 1). While some individuals who have strokes are hypertensive, many are not. Your attempt to estimate stroke is complicated by the widely varying treatments for ischemic, embolic, and hemorrhagic stroke. You state on page 18 that "stroke medications were used to treat a fraction of hypertension cases." Why is this? - none of the medications you count for cerebrovascular disease are anti-hypertensives.

3. According to your methods section, IMS data misses 28% of prescriptions but a proprietary method is used to estimate the remaining missing data. It would help to understand how this method does (or does not) account for the 23% of prescriptions missing by mail order and how this is likely to bias your results.

4. Given the widespread use of antihypertensives and lipid-lowering medications, your estimates of filled prescriptions at the county level appear much too low. If these are accurate, it would help to contrast them with other available estimates and discuss why they differ so dramatically.

Minor Essential Revisions:
1. You refer to figure 1b on page 12 but have no figure labelled as such.

2. Many physicians provide 3 months of medications at a time, and many individuals fill their prescriptions once every 3 months, but you assume that each prescription is for 30 days (since you divide by 12 months). Other individuals fill their prescriptions for only part of each year. How do you account for this?

3. Your results and discussions sections need to be separated.

Discretionary Revisions:
1. In their background, the authors mention NHIS and BRFSS as sources of nationwide prevalence data but do not mention other well-known sources: NHANES, CMS/Medicare data, and National Hospital Discharge Survey.

2. I am concerned that differences between your maps and BRFSS maps are not "suggestive" , as you state on page 19, as much as a reflection of problems with your methods. I fear that you may be comparing very different patient populations.

**Level of interest:** An article of limited interest

**Quality of written English:** Needs some language corrections before being published

**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.