Reviewer's report

Title: Positive Predictive Value of a Case Definition for Diabetes Mellitus Using Automated Administrative Health Data in Children and Youth Exposed to Antipsychotic Drugs or Control Medications: a Tennessee Medicaid Study

Version: 1 Date: 26 March 2012

Reviewer: Eric Weinhandl

Reviewer's report:

Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

- The first paragraph of the background is entirely reasonable. However, it fails to alert the audience to the fact that this study should be interpreted in the context of patients with psychotropic medication use. The authors might improve the introduction by eschewing the discussion of diabetes as a broader public health problem and instead focusing on the putative association between psychotropic medication use (or even medication use, regardless of class) and risk of new-onset diabetes.

- The results would benefit from a simple declaration of how many of the 46 met the computer case definitions of type 1 and 2 diabetes. I can see from the table the answers are 15 and 31, respectively, but I need to review the table to know this.

- The authors could note that 5 of 30 is an estimate of the negative predictive value.

- The introduction could clearly telegraph that the authors have assessed an algorithm for new-onset diabetes vs. existing diabetes. This is not entirely clear until later.

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

- “Cohort membership required at least one year of prior enrollment…” I would assume that this criterion was satisfied with respect to the index date. The authors should confirm this detail.

- “… assured availability” Please change to “ensured availability.”

Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

- As a pharmacoepidemiologist, I am usually concerned about the completeness of prescription drug databases. Prior to the advent of Medicare Part D in 2006, there was tremendous variability in the quality of prescription coverage across state Medicaid programs. I cannot pretend to be familiar with the nuances of Tennessee Medicaid, but I do know that there was a tendency among Medicaid
programs in the South to put strict limits on the number of concurrent prescriptions that would be reimbursed. The authors should seriously, albeit briefly address the historical structure of drug coverage in Tennessee Medicaid, so that readers may have a sense of whether the database at hand included all medication fills vs. only some, due to patients filling prescriptions not billed to Medicaid (because they were paid out-of-pocket or reimbursed through a state pharmacy assistance program).

- “The cohort excluded persons with… medical care indicating diabetes…” This is entirely appropriate, but lacks salient detail. What exactly did the authors exclude with application of this criterion?

“Recent initiators filled a qualifying prescription for a study drug, had no fill more than 90 days prior to the qualifying prescription, and had at least 365 prior consecutive days with no filled prescription for the study drug.” This is difficult to understand and should be rephrased. I think that the authors are making a distinction between use of any psychotropic medication (in the second clause) and refills of the index medication (in the third clause). The audience needs better clarity.

- Why was the adjudication rate so different in the two groups (met case definition [N = 64] and did not meet case definition [N = 113]), if the primary reason for non-adjudication was the same? I am concerned about whether the authors actually randomly adjudicated persons not meeting the case definition, as this has important implications for the sensitivity calculation later in the Results.

- As long as you have calculated sensitivity, you should calculate specificity for the audience, too. Also on the front of sensitivity and specificity, it is not clear to me that the formula for sensitivity is well-placed in the Results. I think that it should either be included in the Methods or omitted entirely.

- Why should I believe the distinction between primary and secondary diagnosis codes in hospital encounters is at all important? Forgive my ignorance, but does Tennessee Medicaid use a prospective payment system based upon DRGs? If so, placement of diagnosis codes may be driven by financial reasons vs. clinical reasons. In other words, the ordering of codes may be clinically irrelevant.

- Applicability of findings to other patient populations is of great concern to me. It may seem rather silly, but I would reorder the limitations. While the algorithm seems plenty applicable to broader patient groups, the authors should be clear that the PPV and NPV estimates here may not be readily applied to non-psychiatric study applications.

- Most importantly, there are no confidence intervals whatsoever in this entire manuscript. This is a small study and it should be made clear (by data presentation) that there is inherent stochastic error in the PPV and NPV estimates that have been presented.

**Level of interest:** An article whose findings are important to those with closely related research interests.
Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.