Reviewer's report

Title: Co-morbidity and health care utilisation five years prior to diagnosis for depression A register-based study in a Swedish population

Version: 1 Date: 14 March 2011

Reviewer: Upali W. Jayasinghe

Reviewer's report:

This manuscript describes a study that is interesting and of significant importance. However, there are some issues that the authors should address.

Statistical review

Major Compulsory Revisions

1. In the abstract, background and elsewhere authors state that during the five-year prior to a depression diagnosis there was a substantially increased level of health care utilisation. Again in the abstract and elsewhere they say that “Patients with a depression diagnosis, adjusted for age, gender and socioeconomic status, used almost twice as many health care resources (physician visits, in-patient care and total number of visits including paramedical staff) during the whole five-year period compared to the controls” and also “There was a strong increase in physician visits one year before diagnosis. The above statements were not tested in multivariate analysis (Table 2) because authors have not included physician visits in both 2001 and 2005 simultaneously in a single model (notes to Table 2 page 19). What authors have tested is whether physician visits in 2001 OR physician visits in 2005 predicted the probability of depression diagnosis. For example; say the number of physician visits in 2001 for a patient was 10 and the number of physician visits for the same patient was 12 in 2005 and the number of visits in 2001 in comparison to that of 2005 is actually lower and the multilevel analysis conducted does not capture this effect. Authors should include both physician visits in 2001 and physician visits in 2005 as independent variables simultaneously in a single analysis with other covariates. A model such as this will show the effect of physician visits in 2005 after adjustment for physician visits in 2001 and vice versa. According to authors Table 2 shows the results of two multivariate analyses (notes to Table 2 page 19) which is not clear because OR of gender, age and socioeconomic status with physician visits in 2001 as an independent variable should be different from those of the model with physician visits with 2005 as an independent variable.

2. Due to the large numbers of subjects in both groups some of the relations may be statistically significant in the logistic regression, although they might not have been in a study with relatively smaller sample (type 1 error). Therefore, a measure of effect size or variance explained should be adopted to evaluate the practical significance of the effects (for example, the Nagelkerk R2).
3. The dependent (response) variable used in logistic regression is not clear to readers. Authors should explain clearly what 1 and 0 in the response variable is.

4. At the top of page 10 authors say “Co-morbidity was included in the model but was found insignificant when adjusted for the number of physician visits (p>0.05)” which means co-morbidity is not an independent factor after adjustment for other confounding factors.

5. Page 9 and elsewhere it is indicated that the likelihood of receiving a depression diagnosis in 2006 was associated with high or low age. This statement is not correct OR for high (80+) and low (30-39) are not significant in Table 2. The result says the likelihood of depression diagnosis in 2006 was lower for age groups 40 to 79 than that of 20-29 age group (reference).

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