Reviewer's report

Title: Population risk factor estimates for abdominal aortic aneurysm from electronic medical records: a case control study

Version: 2
Date: 8 August 2014

Reviewer: Jason Roy

Reviewer's report:

Abstract (& elsewhere): 'protectively associated' could easily be interpreted causally, because 'protective' is causal language. I still prefer positive and negative associations as terminology.

Methods: It wasn't clear to me about when the variables were measured. If we think of date of diagnosis as an index date for cases, was it required that you only used data for that subject on or before the index date? What is the corresponding index date for controls?

Methods: I am a little concerned about the meta-analysis step. My understanding is that variable selection is done separately for each bootstrap sample. That would suggest that for two different samples the selected model might have different variables. If you then want to use a meta-analysis technique to estimate a coefficient, you will be combining estimates from different models. In one model you might be conditioning on 5 variables, and in another you might be conditioning on 8, etc. The coefficients have different interpretations in each model (unless you assume all coefficients of variables not selected are exactly equal to 0). Some clarification here would be great.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.