Reviewer's report

Title: The association of smoking status with healthcare utilisation, productivity loss and resulting costs: results from the population-based KORA F4 study

Version: 4 Date: 11 May 2013

Reviewer: Dennis Petrie

Reviewer's report:

Minor Essential Revisions

The authors have done a very good job at addressing my concerns. My only remaining concern refers to my original comment about the use of the negative binomial model as the second part of the two-part model. My further response is listed below the authors reply to my original comment.

4. Statistical Analysis: The use of the negative binomial regression for “users only” seems slightly strange given that the negative binomial also supports zero use even though by construction the “users” will only have non-zero usage? Some clarification here would be welcomed about why a zero-truncated model was not used or why the negative binomial is more appropriate in this case?

- You are right in that the negative binomial also supports zero-valued observations. However, there are excess zeros in most of our healthcare categories (e.g. for alternative practitioners, where 2,850 out of 3,068 participants did not report health care use). In such situations, the amount of zeros is usually much higher than expected under the count process assumed for the remaining observations with non-zero use. This is why we preferred to model zero counts and non-zero counts separately.

- I understand the excess zero problem and I have no issues with you using a two part model – my concern involves the use of the negative binomial as the second part of the model. In particular you have gone from the case where you had too many zero to the other extreme where you now have too few zeros (none) for the negative binominal model. A regular negative binomial model will try to predict zero counts even though in your case there are no zeros in the second part of your model. Ideally I would have preferred a zero-truncated negative binomial model as the second part of the model. I would be more comfortable if this issue was noted in the text for others looking at following your methods to consider.

Level of interest: An article of importance in its field

Quality of written English: Acceptable
**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

None