**Reviewer's report**

**Title:** The effect of smoking status on healthcare utilisation and costs: results from the population-based KORA F4 study

**Version:** 2  **Date:** 22 January 2013

**Reviewer:** Nadine Berndt

**Reviewer's report:**

This is a well-written and well-designed study. The paper's organization and quality meets the standards of similar manuscripts found in scientific journals. However, there are several points that need to be addressed. I'm certain all of my questions and comments can be revised without problems. I hope they are helpful and instructive, the paper concerns a good topic with significance to the costs of smoking in Europe and should be brought forth to disseminate the findings. Congratulations on completing the project!

**Discretionary Revisions**

1. **Title:** The title only concerns the effect of smoking on healthcare use and costs, however, the authors also looked at indirect costs (such as productivity losses). The title is now misleading and may be adjusted.

2. In the third paragraph of the introduction the authors state ‘the actual impact of smoking on costs of care can be focused on by estimating excess costs of smoking in a data set comprising current utilisation of individuals’. Please reformulate this sentence for enhanced understanding.

3. It is not clear from the introduction which perspective the authors used in their study while it seems that a societal perspective was used (since costs outside the healthcare sector were also measured). This information may be added.

4. Paragraph 4, introduction. What were the a priori identified hypotheses of this trial? Lack of identified primary hypotheses lead to confusion whereas specifying these may enhance the quality of the paper (in the discussion the authors may then refer to the hypotheses).

5. Methods, under data and study design. Were smoking-related diseases not measured in participants (including cardiovascular diseases, respiratory diseases, and cancer)?

6. Paragraph 4, under direct costs (Methods). How did the authors deal with over-the-counter pharmaceuticals? Are those the so called non-pharmacy medications?

7. Results, regression analysis. In Table 3 the authors report ORs for all cofactors, and the 4 groups of smokers. However, only the results for the 4
smoking groups are described in the Results section. What is the additional benefit of presenting the ORs of the other factors? The Tables are relatively large and much information seems to be redundant. Consider shortening Table 3 (and Tables 2 and 4) by only presenting the results of the 4 smoking groups. These Tables could be easily shortened by presenting adjusted ORs for the smoking groups.

8. Tables 2-4. The deviance values and dispersion values are illustrated. Some explanation and elaboration might be included in the manuscript because many readers may be unfamiliar with these model-fit values.

9. Discussion. Some discussion may be incorporated about the fact that smokers may currently have higher costs than non-smokers, but non-smokers live longer and hence can incur more health costs at advanced ages.

10. Discussion. The authors suggest that their findings are in line with other international studies on the costs of smoking and refer to several studies. In California, a study group has done quite a lot of research on the costs of smoking, but a reference to their work is lacking (see for example: Max, W. The Financial Impact of Smoking on Health-related Costs: A Review of the Literature. American Journal of Health Promotion. 2001, 15(5)).

11. Editing by a native speaker might be considered to correct minor misuse of terminology, prepositions, and grammar.

Minor Essential Revisions

1. In the first sentence of the introduction, tobacco smoking should not only be labelled as a health risk, but as a health risk behaviour.

2. Methods. Years of education: 1011 should be adjusted to 10-11 years.

3. Throughout the manuscript the authors confuse the term "participants" with other terms (subjects, patients, individuals, etc). For example in the second and third paragraph under direct costs (Methods), the authors confuse the terms participants with users and patients. Please be consistent in the terminology.

4. Table 1, footnote. The sentence seems to be incomplete. It should be added that those 3 subjects with missing data on school education were excluded from the analysis.

Major Compulsory Revisions

1. Introduction, paragraph 1. The authors suggest that morbidity and mortality associated with smoking has financial consequences for healthcare systems and economies. Evidence for this statement is needed.

2. Introduction, paragraph 3. Information about the smoking prevalence in Germany is lacking. A comparison with the financial burden of smoking in another (European) country should be added since the information about €17.4
billion to €33.6 billion otherwise might be misleading.

3. Introduction, paragraph 4. While the selected bottom-up approach is appropriate and relevant, the introduction section needs to include an improved rationale (including references) for the choice of the bottom-up approach.

4. Methods, data and study design. Although the authors refer to previous studies, some additional information about the design, sampling method and data collection of the S4 and F4 study should be integrated. As such, information about inclusion criteria of participants (if appropriate), and the way the baseline examination and follow-up measurement were conducted is lacking. Were these equally conducted?

5. Methods, data and study design. In other countries than Germany, the education system is different. Please refer to primary, secondary and post-secondary education. Years of education are redundant.

6. Methods, data and study design. Please indicate how smoking status was measured in participants and how it was defined. How much did participants need to have smoked to be treated as a current, occasional former, or never smokers?

7. Methods, data and study design. Distinctions are made between low and risky alcohol consumption, and active versus inactive participants. It is unclear if these distinctions are based upon some national guidelines. If so, please add the relevant references.

8. Methods, direct costs. In the first paragraph more description of the measurement of the costs is needed. Which index year was used, was it 2008? From which year were the national unit costs derived by the AG MEG? Were these unit costs inflated to the year 2008? The study period was from 2006 to 2008, and the follow-up period was longer than 1 year. How were costs and smoking status discounted? Please specify.

9. In the Methods under indirect costs it is stated that the number of days of working absence were restricted to 213. Why 213? In the second paragraph the Friction cost method is mentioned. This approach was originally developed by a group of Dutch economists (see for example Koopmanschap et al., 2005). This reference is lacking and should be inserted.

10. Methods, statistical analysis. How did the authors deal with missing data and outliers in costs (if applicable)? It seems that participants with missing data were excluded from the analyses, but this needs to be specified.

11. Methods, statistical analysis. The way the regression analysis were conducted should be better described. Please clearly explain which approach was used. Was it in one or two steps? As regards Tables 3 and 4, the Odds Ratios of the 4 smoking groups are of real interest, which seem to have been adjusted for the other co-factors. The other information in the Tables seems redundant. Please clarify.
12. Results, Table 1, unadjusted analyses. In describing the characteristics of the study sample, please indicate if the 4 smoking groups were comparable in their characteristics. Regarding the proportions, it seems that current and never smokers might differ on some variables (sex, alcohol consumption).

13. An additional sensitivity analysis might be beneficial especially because of the uncertainty in costs and smoking outcomes that resulted from high risks of recall bias and extrapolation of costs to 12 months. Did the authors consider a sensitivity analysis in which all participants of whom the smoking status could not be specified in F4 /lost to follow-up (28%) are regarded as smokers?

14. Discussion. The comparison to other international research is rather general. The authors mention that their findings are in line with other international studies that have shown increased costs due to smoking. Could the authors give a specific example of an international study that has shown similar findings? Did these studies use a bottom-up approach as well?

15. Discussion. The structure of the discussion seems to be poorly organized, since the authors first give a very short summary of their findings, and then mention their limitations very broadly. I would consider moving paragraphs 11-15 to the beginning of the discussion, this information might be better placed before the limitations. Also, all limitations should be in one section, whereas now they can also be found in various paragraphs (such as paragraph 14). Moreover, the discussion of the main findings is relatively short in comparison to the discussion of the limitations. This needs to be balanced better.

16. Discussion. Please compare the prevalence of smoking as found in your study to the prevalence of smoking in Germany, since this information is now completely lacking. Is smoking prevalence expected to decrease or increase in the future (and what is the expected economic burden)?

17. Discussion. Some costs were measured over the past week or the past 3 months, whereas others were assessed over the past 12 months. As a result, some costs needed to be extrapolated to 12 months and others did not. Recall bias and extrapolation may have led to uncertainty in the estimation of costs. Please give some rationale whether and why these were (un)likely to have affected the validity of the results.

18. Discussion. In the last paragraph before the conclusion, it is stated that another German study using a top-down approach found much less costs per current smoker. Was this difference only due to the nature of the study (top-down vs. bottom-up) or are there also other differences that explain the discrepancy in study outcomes?

19. In the end of their conclusion the authors refer to WHO Framework Convention on Tobacco Control. Can the authors give a specific implication on how to decrease the economic burden of smoking by realising this Framework?
Level of interest: An article of importance in its field

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

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

Declaration of competing interests:

I declare that I have no competing interests.