Reviewer's report

Title: Making use of all available data for monitoring trends in the prevalence of smoking in Portugal: a systematic review

Version: 1 Date: 28 July 2012

Reviewer: Tatiana Andreeva

Reviewer's report:

Review of the article ‘Making use of all available data for monitoring trends in the prevalence of smoking in Portugal: a systematic review'

The presented article is an interesting report making effort to integrate different available tobacco surveillance data and to conclude about the stage of the tobacco epidemic in Portugal. However, several suggestions for major revision of the paper are to be considered in order to improve its quality.

Major Compulsory Revisions

1. First of all, there is no clearly stated research question or research hypotheses authors endeavor to test. As seen from the discussion, one of the research questions was whether smaller sample regional or special groups surveys are reliable for their use in the tobacco use monitoring. Though authors try to conclude on this question in the discussion, they do not directly approach this question in the Results section as well as do not clearly specify methods to achieve that.

2. Second important deficiency is related to selected data analysis. Authors actually do not show how the survey results changed in time. Reader is deprived of the opportunity to see whether there were upward and downward trends in particular spans of time which might be expectable taking into account the specific tobacco control measures implemented in particular years. Such a scatter plot could be helpful in the Results section. Choosing for the linear regression analysis authors actually lose the opportunity to see the variability of the smoking prevalence over time.

3. The third important consideration is related to the measurement issues. With regard to national samples, authors do not comment on how the smoking status was measured. However, in the Table showing results for students it is obvious that just two surveys reported by Steptoe measured current smoking, while all others in fact measured daily smoking. Using both measures in the same regression model leads to the exaggerated range of summary estimates and to the biased estimate of a trend if earlier surveys mostly measured current smoking and later ones measured daily smoking.

4. Another important consideration when longitudinal data is studied is the necessity to distinguish between age, cohort, and historical effects. While both
cohort and historical effects are interesting for the tobacco surveillance and, on the other hand, the existing data may give limited opportunities for that, authors could think about what processes are observed through the acquired data: do same cohorts smoke more/less over time as they become older or do new birth cohorts uptake smoking to a larger/smaller extent?

5. In this study, PubMed is considered the only source of the information regarding the conducted studies while much data can be found only in reports inappropriate for journal databases.

There are several comments for minor and discretionary revisions as well.

6. TITLE is suggested to be reconsidered as it does not point to a significant result or a research question.

7. In the ABSTRACT it is worth showing from which time the studies were searched as it has been mentioned in the Methods section.

8. In the BACKGROUND authors need to better justify why the 'general model' of the tobacco epidemic is insufficient. Though this issue is addressed further, at this point the authors' intentions seem unclear.

9. Instead of the term 'locale-specific' authors might consider using 'country-specific' if this is what was meant.

10. Further on ‘smoking habit’ could be substituted by ‘smoking behavior’ or another conventional term.

11. In the METHODS section, the first paragraph seems not necessary. It distracts attention from the purpose of the submitted paper, makes it sort of supplementary and less important.

12. With regard to use of PubMed, as already pointed to above, there arises a question how searching only PubMed fits with the intention to use all available data. Small regional survey could be less likely published in journal indexed in PubMed. Besides that, there may be data appropriate for the tobacco surveillance as that from EuroBarometer etc.

13. Further authors refer to their search expression; however, it includes many less important details while pointing which keywords were used as the decisive ones seems worthwhile.

14. In the fourth paragraph of the Methods, authors refer to the ‘cardiovascular risk factors under study’ which again points to the larger study and may be recommended to reconsider.

15. In the Data extraction subsection authors refer to the smoking status measurements. However, as I have stated above, caution needs to be taken with regard to the distinction between current and daily smoking, which sometimes is not clearly defined in the original papers but needs to be reexamined by the authors of the systematic review. I also strongly recommend either to limit the
analysis to one of the measures of the smoking status or to consider both in parallel with clear specification of which of the two is considered. It is known that daily smoking (and subsequently current smoking) more readily reacts to effective tobacco control measures, while occasional smoking (which constitutes another part of the current smoking) may be more stable or even change opposite to the changes of daily smoking.

16. Authors devote much attention to the mean age of survey participants' subgroups. If some additional effort is taken with this regard, namely the calculation of the birth year of these same groups, the analysis will allow for distinguishing between age, cohort, and historical effects, which were mentioned above.

17. In the first line on page 8 it is stated: ‘When an age group also included subjects aged below 18 years old (e.g. 17-20 years), we computed the mid-point and excluded the data if the mid-point year was lower than 17.5 years old.’ However, the approaches of the Global Tobacco Surveillance System require that people over 15 years old are considered ‘adult population’ for the purpose of the tobacco epidemic estimates.

18. Data analysis subsection starts with the explanation of the use of linear regression which is a major limitation of the study. While the use of the dummy variables for categorical characteristics (gender, occupational group, etc) does not constitute major problem, assuming linear association with year and age deprives researchers from seeing important variations of the smoking prevalence. Authors may be recommended to switch to at least quadratic estimates if they want to keep the used software approach. To make it simple, continuous variables can be included in the regression models in both their initial and squared form. However, other approaches are applicable depending on the authors’ expertise in statistics.

19. Several examples of terms use in this subsection (like ‘prevalences’, ‘predictions’, ‘conduction’ and ‘investigations’ further on) need to be revised.

20. RESULTS section shows much less than could be expected from reading the Methods. At least the exploratory graph for the trends of smoking prevalence needs to be shown along with the table with the coefficients of the regression model(s). The graph will inform the authors about the approaches to further analysis.

21. Occupational groups with lower smoking prevalence probably need to be characterized in the Methods section with regard to their occupational status. Were these health workers? Teachers?

22. The cause of a wide CI for students becomes obvious from the Table, as explained above. However, it is necessary to decide whether to show these results based on a refined research question.

23. In the first paragraph of the DISCUSSION, the authors somewhat show the research question they wanted to answer; however, this should be made clearer
24. In the second paragraph, they refer to the multivariate model, while the model itself is not shown.

25. Further on, speculations on the causes of lower smoking prevalence among occupational groups and volunteers partially look self-evident, partially simplistic. However, whether to add more explanations like stronger pressure of social desirability or potentially more biased wording of questions depends on the authors’ decision with regard to their main research question and hypotheses they test. The whole discussion needs restructuring based on the research question the authors answer in their paper. The overall expected structure of the Discussion needs to include (1) interpretation of the findings from the point of view of research question and advancing the area of research, (2) comparison with other study results, (3) pointing to strengths and weaknesses of the current study, (4) suggesting further advancement of the research area. The authors attempt to address several research questions makes it difficult to keep to this traditional structure.

26. In the second part of the DISCUSSION, authors approach the issue of the tobacco epidemic development in Portugal which seems a valuable outcome of the effort. However, based on the linear regression analysis it is impossible to explore whether any ‘peaks’ have been reached or overcome. Besides that, authors might be willing to take into account that the well-known descriptive model of the tobacco epidemic was based on the ‘observational data’ from most developed countries; in other words it described what happened with smoking prevalence and mortality before the tobacco control measures were widely implemented. Currently, with the FCTC in place, in many countries the course of the tobacco epidemic has changed dramatically. Portugal is not an exception. To be more specific, according to the European Commission data http://ec.europa.eu/taxation_customs/taxation/excise_duties/tobacco_products/rates/index_en.htm annual cigarette sales in Portugal in early 2000s were about 18 billion cigarettes, and in late 2000s they decreased to 12 billion. The most probable cause for sales decline was the increase in cigarette taxes and prices: the most sharp price increase was observed in 2005-2008: from 2.4 to 3.3 euro per pack. Series of surveys, conducted in 2006, 2007 and 2010, demonstrated that current smoking prevalence in Portuguese adults decreased in 2006-2010: from 30% to 28% in men and from 27% to 23% in women. According to the Eurobarometer surveys, current smoking prevalence in Portugal declined from 29% in 2002 to 27% in 2005, and then to 24% in 2006 and 23% in 2009. Authors might be willing to consider their data from this point of view.

27. The first sentence in the CONCLUSION does not seem balanced with the results and if authors decide to keep it, it needs to be reflected in earlier sections of the paper.

Overall, the paper is promising to be helpful for the development of country-specific tobacco surveillance systems, and authors are wished every success in their effort on this way.
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