Author's response to reviews

Title: Methodologies used to estimate tobacco-attributable mortality. A review.

Authors:

Monica Perez-Rios (mrmonic@usc.es)
Agustin Montes (mrmonte@usc.es)

Version: 3 Date: 29 June 2007

Author's response to reviews: see over
Dear Editor,

We have reviewed the article entitled "Methodologies used to estimate tobacco-attributable mortality: A review" and we are sending you the manuscript taking into consideration the reviewers’ commentaries. In our opinion their comments have helped to improve the quality of the manuscript. In addition, three native English speakers have reviewed and edited the paper.

You can find below a point-by-point response to the reviewers' reports, detailing where and how we have revised our manuscript in response to the reviewers' comments.

The author responsible for correspondence is: Dr Mónica Pérez Ríos. Department of Preventive Medicine and Public Health. University of Santiago de Compostela. Santiago de Compostela. CP 15704. Spain. Phone: 34-981-540044. Fax: 34-981-572282. mrmonica@usc.es

The manuscript has been read and approved by the two authors.

Thank you for considering the manuscript and we will wait for the future decision. If you or the reviewers need more information, do not hesitate to contact us.

Looking forward to hearing from you,

MPR AMM

Mónica Pérez Ríos and Agustín Montes Martínez.
The CDC’s SAMMEC (Smoking-Attributable Mortality, Morbidity, and Economic Cost) computer software application (58) uses this methodology. SAMMEC is a software package commonly used in the United States to estimate attributable mortality due to smoking, years of potential life lost and indirect mortality costs. SAMMEC computes PAF automatically after the user includes prevalence of tobacco consumption. Furthermore, the user must supply the number of deaths by 5-years age groups from 35 or older, for each smoking-related diagnosis. Estimations from SAMMEC can include attributed deaths to fires and secondhand smoke. The Simsmoke model, a model that predicts the effect of policies on smoking rates and deaths attributable to smoking, uses this computer application to estimate deaths attributable to smoking.

Cancer deaths due to smoking are calculated as the difference between observed and expected deaths in a population. To apply it, age- and sex-specific cancer mortality rates are needed, and age- and sex-specific cancer mortality rates for non-smokers are computed on the basis of the CPS study. The expected deaths are related to the number of deaths that would occur if the whole population was formed by non-smokers. To calculate the expected number of deaths the follow-up, over 12 years, of the never smokers enrolled at the CPS I study were employed and death rates for cancer were computed. These rates were applied to the estimated number of person-years of exposure for non-smokers to obtain the expected number of deaths for each cancer. The attributable fractions calculated have been found to be similar to those yielded by the CPS.

The method proposed by Rogers and cols. combines prevalence and mortality risk rates in order to offer more precise estimates of smoking attributable mortality. This calculation procedure tries to avoid some problems related to previous methods as the use of risks derived from non-random sample of population, risks rarely adjust for confounding factors or the classification of the smoking status in crude categories without attending to the number of cigarettes smoked by former and current smokers. At first, age-specific smoking prevalence and mortality risks were estimated. The authors define 7 population groups distinguished by reference to the amount of cigarettes smoked, and classifies them by sex and age-group (p): non-smokers, light smokers, moderate and heavy smokers, light ex-smokers, and moderate and heavy ex-smokers. To determine the risk of death due to cigarette smoking, discrete time hazard models were employed. Roger and cols matched data of a health interview survey to mortality data. The exponentiated values of the logistic regression coefficients yielded can be interpreted as odds ratios.

The next step is to determine how many people would be in each smoking status (n):

\[ n = p \times Pop \]

being Pop the age-sex-specific population in the area studied.

The last step is to estimate the excess risk of death (R) of each smoking status relative to never smokers:

\[ R = mx, c - mx, n \]

where \( mx, c \) is the age-specific central death rate for each smoking status and \( mx, n \) is the age-specific central death rate relative to never smokers.
Finally the excess number of deaths is calculated as follows:

We agree with the reviewer advice and the following paragraph has been included:

The calculation procedure was described in detail in a thesis and summarized elsewhere. Owing to the reduced scale on which the method has been used their calculation procedure is not described in the present paper, but it is important to introduce two epidemiological effect measures applied in it: the “potential impact fraction” and the “trend impact fraction”. Both are indicators of the reduction in the incidence of a disease in the population studied, the former reflects changes in the evolution of a disease after an intervention and the latter is referred to autonomous or natural trends.
Reviewer 2

1st paragraph. I assume only 3 studies used 'individual analysis', ie those referenced? Might pay to be explicit.

We are in agreement with the warning of the reviewer. The sentence was changed as it follows:

“The first is based on individual analysis of deaths to ascertain what, if any, implication tobacco use may have had in them. Only three studies have applied this procedure.”

The last paragraph of Discussion is distracting - nowhere previously in paper was genetics mentioned. It reads as an afterthought at the moment.

We are in agreement with the advice of the reviewer and the sentence has been deleted.

We have read with great interest the suggested papers. Nevertheless, they have not been cited because they fall outside of the inclusion period of the search literature, which finishes in 2005.
Reviewer 3

The key weakness is that the authors don’t illustrate any meaningful differences in methods of estimating tobacco deaths. A table that compares lets say the US middle age male deaths via these methods would inform the reader considerably.

Although this is an important and essential point, this is not the paper objective. As we explained in the text different methodologies need different information and some methods are applied to certain causes of death and some to others. Also the eligibility of a population (the reviewer’s proposal is the US middle age male) can modify the output.

Although we think that it would be better not to include these results, we have done a table comparing the output of four of the methods explained by applying them in the Galician population, the cause of death studied was lung cancer, and the results are as follows:

<table>
<thead>
<tr>
<th></th>
<th>PAM Cohort study</th>
<th>PAM Case control study</th>
<th>Peto’s method</th>
<th>Garfinkel</th>
</tr>
</thead>
<tbody>
<tr>
<td>35-64 years</td>
<td>357</td>
<td>272</td>
<td>333</td>
<td>339</td>
</tr>
<tr>
<td>65 and over</td>
<td>726</td>
<td>343</td>
<td>747</td>
<td>731</td>
</tr>
</tbody>
</table>

As this is not the objective of the paper and also it is now too long, we have decided not to show these results.

We agree with the reviewer’s comment and the sentence was excluded.
We are in agreement with the reviewer’s comment and “roughly accurate” was changed by the term reliable.

The Methods Section is far too sparse. The authors need to be much more explicit and detailed about the search strategy and the criteria used to discard the majority of the 372 papers retrieved.

We are in agreement with the reviewer’s comment and we have added:
“The use of an epidemiological method to estimate mortality attributable was the main inclusion criteria. Mortality descriptive papers, as cohort follow-ups or mortality studies were not included unless an epidemiological analysis has been done.”

When the search strategy was planned the preference of an exhaustive one was preferred in order to avoid the exclusion of important information.

Although this was not included in the draft, more information about the papers excluded can be provided:

373 papers retrieved: 19.9% explain the method applied
- 8.2% are based on autopsies
- 9.3% description of death certificates
- 11.9% uses mortality databases without death attribution
- 4.4% follow-up studies
- 46.3% do not explain the procedure applied or do not apply any procedure at all.

The papers included in the methodological review were not sponsored by tobacco industry. Perhaps authors that have published papers or letters criticizing the methodological papers could be supported by tobacco industry, as the one cited by the referee. This author included in the acknowledgments of the paper entitled “Mortality from tobacco in developed countries: are indirect estimates reliable?” the British- American Tobacco, Imperial Tobacco, and Rothmans; for this reason, as the reviewer suggested, these relations were included in the text as follows:

“This method has not been exempt from criticism directed mainly at the calculation of summarized prevalence. Some of these critics were supported by the tobacco industry. “

We appreciate reviewer’s comment about Peter Lee, and the initials were changed.

Page 5: How were the papers which dealt with “mortality attributable to different causes” accessed; there is no mention of this.

We agree with the reviewer’s comment and a paragraph was added in the methods’ section and in the results’ section:

Methods:
“In addition to tobacco, the estimation of attributable mortality is also applied to other risks factors as alcohol consumption or obesity. In order to avoid the exclusion of valid methodologies that can be used to estimate tobacco attributable mortality the same search was done excluding the terms tobacco or smoke”

Results
“The search without the terms tobacco or smoke did not identify new methodologies.”

We agree with the reviewer comment and the sub-heads terminology was changed in some methods. The “hybrid method” term was changed. Instead of this we decided to distinguish between the methodologies applied with relative risks derived from case-control or cohort studies, so the sentence was modified into:

“Due to the methodology applied it is possible to join them into four categories: Prevalence-based analysis (Prevalence-based analysis in cohort studies, Prevalence-based analysis in case-control studies and the basic model), Peto and colleagues’ method, methodologies based on the calculation of excess mortality (Garfinkel’s and Roger’s method) and predictive models (Prevent).”

We agree with the reviewer and this section was changed, especially the Garfinkel method, the Roger method and the predictive models. The text was changed as follows:

**Garfinkel’s method**

Garfinkel’s method has been applied to estimate cancer mortality attributable to tobacco use. Cancer deaths due to smoking are calculated as the difference between observed and expected deaths in a population. To apply it, age- and sex-specific cancer mortality rates are needed, and age- and sex-specific cancer mortality rates for non-smokers are computed on the basis of the CPS study. The expected deaths are related to the number of deaths that would occur if the whole population was formed by non-smokers. To calculate the expected number of deaths the follow-up over 12 years of the never smokers enrolled at the CPS I study were employed and death rates for cancer were computed. These rates were applied to the estimated number of person-years of exposure for non-smokers to obtain the expected number of deaths for each site cancer. The attributable fractions calculated in this way have been found to be similar to those yielded by the CPS.

**Roger’s method**

The method proposed by Rogers and cols. combines prevalence and mortality risk rates in order to offer more precise estimates of smoking attributable mortality. This calculation procedure tries to avoid some problems related to previous methods as the use of risks derived from non-random sample of population, risks rarely adjust for confounding factors or the classification of the smoking status in crude categories without attending to the number of cigarettes smoked by former and current smokers. At first, age-specific smoking prevalence and mortality risks were estimated. The authors defined 7 population groups distinguished by reference to the amount of cigarettes smoked (p) and classified by sex and age-group: non-smokers, light smokers, moderate and heavy smokers, light ex-smokers, and moderate and heavy ex-smokers. To determine the risk of death due to cigarette smoking, discrete time hazard models were employed. Roger and cols matched data of a health interview survey to mortality data. The exponentiated values of the logistic regression coefficients yielded can be interpreted as odds ratios.

The next step is to determine how many people would be in each smoking status:

\[ n = p \times \text{Pop} \]

being \( n \) the age-specific population in the area studied.

The last step is to estimate the excess risk of death of each smoking status relative to never smokers:

\[ m_{c} - m_{n} \]

where \( m_{c} \) is the age-specific central death rate for each smoking status and \( m_{n} \) is the age-specific central death rate relative to never smokers.

Finally the excess number of deaths is calculated as follows:

\[ \text{ED} = \]
Predictive models

Under this head, one method was in evidence, namely: the Prevent model. Owing to the reduced scale on which this has been used or their use for more specific cases, their calculation process is not as well known. The Prevent simulation model was developed in 1988 in The Netherlands and is regarded as being a multifactorial generalization of the etiologic fraction. It has basically been used to predict mortality due to various causes, including tobacco (89). The methodology used allows, among other factors, for a temporal dimension to be considered and envisages the possibility of a risk factor being associated with more than one disease and a disease being associated with more than one risk factor, i.e., multiplicity of cause or effect. The process of calculation is laborious and calls for knowledge of multiple data, such as birth- and mortality-rate series or the likelihood of dying at different ages for each sex. The calculation procedure was described in detail in a thesis and summarized elsewhere. Owing to the reduced scale on which the method has been used their calculation procedure is not described in the present paper, but it is important to introduce two epidemiological effect measures applied in it: the “potential impact fraction” and the “trend impact fraction”. Both are indicators of the reduction in the incidence of a disease in the population studied, the former reflects changes in the evolution of a disease after an intervention and the latter is referred to autonomous or natural trends.

We agree with the reviewer’s comment. In order to avoid this problem the paragraph was excluded and the text was modified in the first method as follows:

“Attributable deaths are calculated for each cause of mortality, using the following formula:
AM = OM * PAF; where AM is mortality attributed to tobacco, OM observed mortality, and PAF the population attributable fraction.
To ascertain the PAF different methods of calculation exist, though the most widely used in this particular instance, is based on the formula proposed by Levin”

As pointed out elsewhere by the authors the PAF is a function of RR, exposure prevalence (at a given point in period) and of the maturity of the epidemic. So for example PAF for lung cancer is 90% and in Asia about 70% at present.

We agree with the reviewer comment.

Lam and McGhee applied in the estimation of attributable mortality RR obtained as OR after the development of a case-control study, and consequently the formulae to calculate the PAF is different to the one applied by the SAMMEC.

Incidently although McGhee calculated PAF these were not quoted in reference 80 (only RR were presented) but rather in a further analysis in which the community costs of tobacco were estimated (McGhee et al Tobacco Control 2006; 15:125-30).

Sorry, it was a mistake. The paper was excluded but the new one was not included because it was published after the search was done.
Wan agree with the referee comment, but we can not include this study because an explicit attribution method was not applied.

Page 9: Edit out abbreviations.

P0, p1, p2, RR1, and RR2 explanations were excluded

Page 10: English syntax a big problem here. The text also goes from specific (Peto, Lam, McGhee) to GENERAL (Gerfinkel) to VAGUE (Rogers). The text could be made much clearer and more focussed and applied.

The text was modified all along the manuscript where necessary. The modified text was located in pages 2 and 3 of this answer to the reviewer.

Page 15: I am not sure what they mean with the statement about “the dose-response gradient”. There is relatively little variation in amount consumed among regular smokers and even one cigarette/day is associated with significant mortality risks. Average daily consumption among regular adult smokers is usually high but may vary between countries. Of course data structure can always potentially be refined, but this is a different issue from a critique of the basic methodologies.

The reviewer’s explanation about the response-dose gradient is correct, but perhaps our sentence is confusing. We have changed the sentence as follows:

“To view smokers as a single entity could lead to a distorted mortality estimate, since failure to take account of the number of cigarettes smoked, age at initiation, years smoking and other variables that could modify risks values.”

Page 16: One of the important points about the Lam/McGhee methodology is that it was based on ascertainment of active smoking/passive smoking 10 years before death and so avoided bias which may arise because sick people quit in the years closer to death. This is also relevant to the comment on PAGE 5 about studying tobacco use “in the studied populations prior to the start of the estimations”. However I think re-classifying quitters as non-smokers is invalid in view of the rate at which risk is shed and the residual risks associated with long term smoking.

Perhaps the sentence sense is not clear. The classification of the population in relation with their tobacco consumption is mainly done by themselves. We did not want to say that the researchers working on smoking attributable mortality must reclassify former smokers; we tried to say that life-long former smokers reclassify themselves as never-smokers.

Page 19: I do not accept that “Mendelian randomisation” has anything practically useful to offer the public health assessment of the impact of tobacco on mortality. Where smoking prevalence declines, so do population age-specific death rates from lung and other cancers and other tobacco-induced diseases. While accepting that individuals are more or less susceptible because of their genetic make up, in the bigger picture these populations have not changed their genes, only their exposures. For tobacco-induced disease, the biggest single preventable cause of mortality in most countries, valid estimates of tobacco mortality are important drivers of public health policy and rapid achievement of the goals of the WHO Framework Convention on Tobacco control.

The paragraph has been deleted.