Author's response to reviews

Title: Indirectly estimated absolute lung cancer mortality rates by smoking status and histological type based on a systematic review

Authors:

Peter N Lee (PeterLee@pnlee.co.uk)
Barbara A Forey (BarbaraForey@pnlee.co.uk)

Version: 2 Date: 5 March 2013

Author's response to reviews: see over
Indirectly estimated absolute lung cancer mortality rates by smoking status and histological type based on a systematic review

Point-by-point reply to Reviewers Comments

Peter N Lee and Barbara A Forey

Date: 5 March 2013

General comment
We thank both referees for their kind words, their helpful comments, and the time they have taken on reviewing our paper. Our replies to the points made are shown in bold in the text below.
Reviewer: Joshua Muscat

Reviewer's report:

This review is a monumental undertaking and a significant contribution to the literature. The study compiles a large bulk of the descriptive epidemiology of lung cancer, which addresses some important ongoing issues such as relative rates in women vs. men and many others.

There are clearly some limitations in the approach and the voluminous data and discussions take this into consideration.

As extensive as this current manuscript is, there are still some slight modifications I feel are necessary.

The abstract should make clear that the reviewed literature is highly selective. A systematic review usually implies the universe of literature but the authors excluded many studies (e.g. occupational etc.). The abstract methods should note this.

**Following the first sentence of the methods**

“We estimated study-specific absolute mortality rates and variances by histology and smoking habit (never/ever/current/former) based on relative risk estimates derived from studies published in the 20th century, coupled with WHO mortality data for age 70-74 for the relevant country and period.”

We have added the sentence
“Studies with populations grossly unrepresentative nationally were excluded.”

There is not space in the abstract for anything more detailed. As it is I have had to make minor rewording changes in other parts of the abstract so as not to exceed the word count. We also removed the word “Typically” from page 10 (last paragraph of the “Application of the method” section), because we have described all the exclusion criteria, not just typical examples. The reasons for study exclusions are shown in the Additional File.

My second comment is that most RR in the review were unadjusted or age-adjusted. That is fine. In ever smokers, some of the findings observed might be due to differences in cigarettes per day. For example, men smoke more CPD than women. The discussion section should comment about whether CPD differences might account for some of the observed findings by sex, country etc. The second paragraph of the discussion under “ever smoking rates” already mentions the fact that differences in amount smoked, and in other aspects of smoking, might contribute to regional differences, and I do not think it necessary to extend this discussion. However I have amended the previous paragraph on differences between the sexes so that it now reads:

“The excess in rates for males is more evident for ever smokers than for never smokers. This is unsurprising in view of the higher prevalence of smokers in males, their greater daily cigarette consumption, and their earlier take up of the habit.”

(i.e. adding the phrase “their greater daily cigarette consumption”).
Reviewer: Stephen Morrell

Reviewer's report:

General

This is a useful but long paper that can be reduced considerably by relegating many of the tables and figures to an Appendix (online). Concise summary tables covering each of the 4 major domains (ever vs. never; current vs. ex. vs. never; squamous vs. adeno; metaregression) are needed that better reflect the description of results in the text.

When publishing in an online journal, with no restriction on length, it seems easier to have all the material in the paper itself except for technical information only the very occasional reader will want to access. We could relegate Tables 4 to 9, which present the study-specific indirect estimates, and perhaps Table 3, which gives study details, to Additional files, but prefer not to as we feel they help the reader to gain insight into the data. Other tables seem essential to stay. We could relegate all the Figures to Additional Files, but we prefer not to as they help the reader (some of whom find Figures much more useful than Tables) to assimilate the results easily. Our preference, therefore, is to leave all the Figures and Tables as they are, but if the journal editor really feels we must move Tables 3 to 9, and/or Figures 1 to 6, to Additional files we will do the necessary work.

Particulars

p.9, 2nd para: The authors state “Relative risks for current and ex smoking (each relative to never smoking) were included only where their definitions (age, race,
product smoked etc) precisely matched each other.”, but fail to justify this tight criterion for inclusion/exclusion. This contrasts with the authors’ much looser criterion for ever- vs. never-smoker analyses, as in “However, there was no constraint that that definition had to match the definition of the ever vs never relative risk,...” The concluding part of the latter sentence makes this rather murky (“..., and separate versions of the 2×2 (never/ever) and 2×3 (never/ex/current) tables were used where they did not match.”). What is excluded or included and what is being matched?

Studies excluded (or partly) with RRs derived from comparison groups that did not match according to the criteria above are not listed in the file of excluded studies, nor is non-matching on the above criteria listed as an exclusion category or reason for exclusion indicated in the list. In the absence of justifying these tight versus loose criteria it would be prudent to perform a sensitivity analysis that compares summary estimates for never/ex/current smokers with these studies included versus excluded.

We are sorry if we gave the impression that we had made “murky” exclusions. The text was in fact intended to convey the opposite, that we made efforts to include all relevant data, even when a pair of current/ex RRs were not available that matched the ever RR. It is necessary for the logic of the method that the current and ex smoking RRs must match each other. The paragraph has been deleted and relevant parts of it added as a new second paragraph (shown below) of “Application of the method” section (page 8), which hopefully clarifies what we did. Table 3 indicates where the current/ex RR pair do not
match the ever RR – most commonly by showing “E” on a separate row from “C,X” with a different entry in the “product” column.

“Where there was a choice, relative risks for smoking of any product were selected if available, or of cigarettes (or cigarettes only) if not, then selecting the widest available age and race group, and, for prospective studies, the longest follow-up. Current and ex smoking relative risks were constrained to match each other on these selection criteria, but not necessarily to match the ever smoking relative risk. Where relevant (e.g. when using relative risks for ever smoking any product and for current and ex cigarette smoking) separate versions of the 2×2 (never/ever) and 2×3 (never/ex/current) tables were used, and the indirect estimate of the never smoker rate that is reported is that based on the never/ever comparison.”

There were no studies wholly excluded in this respect, so no change has been made to the Additional File. There were a very few studies providing current and/or ex smoking relative risks that could not be used, and these are now indicated by an extension to footnote g of Table 3

“g E = ever vs never, C = Current vs never, X = Ex vs never. Studies with no ever vs never relative risk were excluded (see Additional File). Except where indicated below by footnotes l-o, studies shown only as “E” had no current vs never or ex vs never relative risk”

and additional footnotes l to o

USA : HAMMON and UK : ALDERS
“l Current smoking excluded because no ex smoking relative risk available”

UK : DARBY

“m Current and ex smoking excluded because no matching pair of relative risks available”

W Europe : VUTUC

“n Current and ex smoking excluded because only available relative risks did not satisfy age criteria”

E Europe : JEDRYC

“o Ex smoking excluded because no current smoking relative risk available”

Sensitivity analysis is not possible, as one cannot calculate the indirect estimates for current smoking in the absence of corresponding estimates for ex smokers (and vice versa).

p.11, 2nd para: The authors chose the age group 70-74 years as the best fit between observed (from CPS-I) and modelled estimates of Lo. However, from Table 2 the best fit in males was 75-79 years while the best fit in females was 70-74 years. Why not use each of these age groups separately for each sex, since separate estimates by sex overall are being produced?

The main reason against using age 75-79 years for males and age 70-74 years for females is that we could not then have validly compared the results for the two sexes, which was one of the particular interests. Also it would have made the paper more complex. As we say on page 11, the correspondence overall
between the observed and predicted rates was best for age 70-74 so we used that.

p.12, 2nd para: Interval 3 (1971-86) overlaps with interval 4 (1981-90), which appears to be a typo.

**Typo corrected**

pp.16-17, section ‘Trends in rates for never and ever smokers by region’: this section is not needed as descriptions of figures repeat previous descriptions. Anything new here can be added to previous description with figures cited there. **We prefer to leave this relatively short section in. The figures present results for never smokers and ever smokers in the same graph, and it would lead to a lack of clarity if one tried to refer to the Figures separately in the earlier separate sections “Never smokers” and “Ever smokers”**.

p.20, last line: change ‘causes’ to ‘caused’

**Typo corrected**

p.21, 3rd para, 4th line: replace period with comma

**Typo corrected**

p.26, 1st para, lines 3-4: Delete ‘Bias due to misclassification of smoking status is also relevant.’, as this issue is taken up on pp.27-28, otherwise it appears as if it’s asserted without references or its likely effects on estimates.
The paragraph has been slightly rewritten to make it clearer which are the various limitations which will be discussed in the following paragraphs. It now reads:

“There are also a number of limitations with the data on relative risk by smoking habit obtained from the IESLC database. These include variations in definition of smoking, definition of disease and extent of adjustment for confounders, and bias due to misclassification of smoking status. These and some other issues are also discussed in the first paper on IESLC [3], but some of the principal points are considered below.”

p.28, 2nd para: This example of mis-classification is misleading. If 2.5% of ever smokers in the population are mis-classified as never smokers, then the true prevalence of ever smokers in a population with 50% ostensible ever smokers is 51.28%, not 48.75%. (ie, solving for x in x-.025x=50). Consequently, the estimates are biased toward the null (ie, lower than in reality), not away from the null as suggested by the authors.

The paragraph is correct and has not been changed. The start of the paragraph makes it clear that the 50% is the number of true ever smokers. Misclassification of 2.5% of these as never smokers means that the observed number of ever smokers will go down to 48.75%, the observed number of never smokers will go up to 51.25%, the observed relative risk will be less than it really is, and the calculated rate in never smokers (based on the observed data) will be too high. I have, however, changed the last sentence of the previous paragraph so that it reads
“If these rates apply, and there are considerable uncertainties [39,44],
misclassification will have a marked effect on the estimated lung cancer
death rates in never smokers.”
(deleting “and ever smokers” at the end)