Reviewer’s report

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

Version: 1 Date: 22 February 2013

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.

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.

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?

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

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.
p.20, last line: change ‘causes’ to ‘caused’
p.21, 3rd para, 4th line: replace period with comma
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.
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.