Reviewer’s report

Title: Lung cancer mortality in towns near paper, pulp and board industries in Spain

Version: 2 Date: 14 May 2008

Reviewer: Mark Clements

Reviewer’s report:

This is a very interesting paper which is both well written and well presented.

Minor Essential Revisions

1. I have two concerns with the epidemiology and biostatistics of the paper. First, the authors obliquely note (first paragraph, page 9) that there are concerns with “the latency of lung cancer”; this is a critical point, where the authors have made the strong assumption that exposure today will tell us something about exposure over the last 20-30 years. It would be useful for the authors to expand on this point. In particular, is there any evidence as to whether air pollution works as an initiator or as a promoter for lung cancer?

2. Second, the authors do not discuss the issue of multiple comparisons, particularly when interpreting associations for each industry site (e.g. Table 3). In the light of multiple comparisons, the authors should interpret the association for Industry 3730 very carefully; the association for Industry 2761 would probably be robust, even given a conservative Bonferroni correction.

3. Abstract, Conclusions, first sentence: “_a_ useful” [phonetic]

4. Background, third paragraph: could the authors summarise the grade of evidence for the associations in the paper, pulp and board industries, please?

5. Background, second to last paragraph, last sentence: what will be the effect of voluntary reporting on representativeness?

6. Methods, paragraph 2: Under ICD-9, code 162 is “Malignant neoplasm of trachea, bronchus, and lung”; under ICD-10, C34 is “Malignant neoplasm of bronchus and lung” and C33 is “Malignant neoplasm of trachea”. This suggests that both C33 and C34 should be included for comparability with code 162.

7. Methods, paragraph 2: would there be any differences between populations from the “1996 municipal roll and 2001 census”?

8. Methods, second paragraph on page 5: what happened to the “intermediate group”? Were they excluded from further analysis? It is questionable whether their risk will be “intermediate”, as other sources of air pollution may be more or less harmful.
9. Methods, third paragraph on page 5, fourth sentence: can you re-phrase the expression “proportion of illiterates”?

10. Methods, third paragraph on page 5: how were the variables for adjustment chosen? How would this relate to potential confounding with smoking status?

11. Methods, fourth paragraph on page 5, second sentence: how many industries had neighbouring towns (ie how many industries were included in the analysis)?

12. Methods, last paragraph: how were cases and person-time calculated for the concentric bands?

13. Results, first and second paragraphs: could the authors provide some commentary on Tables 1 and 2, please?

14. Results, fifth paragraph: I suggest care in the interpretation of site 3730, given multiple comparisons.

15. Discussion, second paragraph: are the authors suggesting that occupational case-control and cohort studies can only be used to “formulate hypotheses and guide research”? I agree that other study designs would provide further evidence on such associations, however I suggest that the occupational case-control and cohort studies will remain the gold (or silver?) standard for assessing such associations.

16. Figure 3: How were the curves estimated?

Discretionary revisions

17. There are a large number of references (n=68) for such a concise report. It would be useful if the authors could be more selective in their referencing.

18. Background, first paragraph, last line: “to five years after diagnosis”?

19. Background, second paragraph, first sentence: “which corresponds to the diversity”?

20. Background, second paragraph: when the authors discuss “principal aetiological agent[s]”, the agents are said to “contribute to” or are “related to” x% of disease. Assuming that are authors are referring to population attributable fractions, the authors could use the less ambiguous expression “attributed to”.

21. Background, third paragraph, first sentence: “its workers are exposed to a range of toxic agents [20-24] which have been linked to lung cancer...”

22. Methods, paragraph 2: please specify the 18 age groups (I assume that these are 0-4 years, 5-9 years, ..., 80-84 years, 85 years and over).

23. Methods, first paragraph on page 5, last sentence: the cited reference to verification and correction using a GIS is valuable; out of interest, how common were such corrections?
24. Methods, third paragraph on page 5, first sentence: “with _the number of expected cases used as an offset._”

25. Methods, third paragraph on page 5, second sentence: this sensitivity analysis is well conceived.

26. Methods, last paragraph, last sentence: the use of Dean’s test is appreciated.

27. Figures 1-2: these are elegant figures. Following the use of forest plots from meta-analyses, is there any chance that each estimate could include a box whose width is inversely proportional to the variance of the estimate?

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.