Reviewer’s report

Title: Lung cancer mortality and exposure to polycyclic aromatic hydrocarbons in British coke oven workers

Version: 2 Date: 21 June 2013

Reviewer: GRAHAM GIBBS

Reviewer’s report:

This is an interesting and important paper and considerable effort has gone into the analyses. However, the Hurley paper on whose study this analysis was published in 1983 and while the IOM reports are on the WEB, the 20 year follow-up by Hurley et al has not been published in the conventionally peer reviewed literature. Therefore, in my view, this paper is incomplete without the reader being able to examine the quality of the data used in these statistical analyses.

The literature from Hurley’s time needs updating. There is mention of some papers, such as that of Armstrong et al, but not a systematic review of this topic. IARC are mentioned, but only for their classification. The Constantino paper in 1995 is mentioned. If these are the only papers on coke oven PAH exposure-response relationships, then the article should say so or reference any other studies of coke workers. Most are the studies mentioned in the Armstrong paper.

More importantly the epidemiology needs to be described in detail, as the reader, without researching the IOM report cannot evaluate to quality of the data used in the analyses. What was the original size of the cohort, who and how many were excluded and why? There were workers who were lost to follow-up. How many workers were lost to follow-up? How did they differ from those followed completely? The original study by Hurley noted that the vital status of 71 NSF and 82 of the BSC men were undetermined. How did you deal with this? Were they assigned a cause in proportion to the rest of the workers? What was their exposure compared to the other workers. Were there differences between the two plants in follow-up? What was the actual size of the cohort as originally followed and then analysed. Table 1 attempts to show this but when one adds the NSF and BSC numbers provided in Table 1 they total 6405 men. In the Hurley paper they report on 3812 NSF workers and 2708 BSC workers – total 6520 who were over 20.5 and 20 years at NSF and BSC respectively. The authors table 1 suggests that there could have been some workers who were less than 20 years of age in 1967, so the difference would be greater. Why the apparent discrepancy? What was the failure rate in follow-up on the analysed group.

The study was based on a cross-sectional definition of a cohort which survived to 1967 when they entered follow-up. They were then followed for 20 years. It is not
clear whether table 1 shows all workers hired or the workers who were in the cohorts. The heading suggests the latter. If this is the case, it appears that the cohort included men born as early 1890 and 753 at NSF and 415 who were 57-67 years of age in 1967. If I am correct for analysis there were 192 men in this age group. To be alive in 1967, they were survivors. As a result, it is not unusual for the first few years of such a cross-sectionally defined cohort to show have a lower rate of a disease such as lung cancer. Do you have data that you can show by decade of follow-up?

A table such as that in the Hurley paper showing the actual levels of BSM and B(a)P by work area would be useful.

There is no mention of the number of measurements on which the exposure estimates were made. They should be included

A discussion of why the B(a)P did not change with time while BSM did would be useful.

A discussion of why there might be such a difference in experience of workers on oven tops at the two companies. Is there any indication that there are between plant differences within the same company? Perhaps numbers are too small to look. It is now 2013. The latest follow-up was 1987, why no follow-up that would add another 25 years? Is that worth suggesting if the data are important for compensation.

There is general acceptance that the arithmetic mean is more appropriate for exposure estimation than averages of geometric means. Why were they not used if available? The impact of this on the relationships should be discussed.

In the conclusion, it is mentioned that the relative risk for the more than 5 years on oven tops was 1.81 whereas Hurley reported a significant 2.10. He found that the less than 5 year workers had a relative risk of 0.94. Your comparable value (0.95) would be helpful to state. In fact the 1.81 is significant and should be mentioned.

Specific comments:

I do not like the term ovens tops, it is awkward language – can you perhaps use “oven tops” defining what you mean by tops and sides?

In the abstract, The last 2 lines of background might read better as “estimates to BSM and to benzo(a)pyrene (B(a)P”. The “also, including…..latency “ can be eliminated as this is repeated in the methods section. Further, the lines in results “.. results for those with 0-5 years…… did not suggest a trend ….strong” is not clear. Did you mean that risks for the 0-5 year workers were not statistically lower than those of the more than 5 year workers, but risks were apparently lower suggesting a possible exposure effect?

Last line of conclusions “real effect” needs clarification. Perhaps “statistically supportable “(???) effect of BSM or B(a)P on lung cancer risk. What does real
mean?