Reviewer’s report

Title: Exposure Measurement Error in PM2.5 Health Effects Studies: A Pooled Analysis of Eight Personal Exposure Validation Studies

Version: 1 Date: 13 October 2013

Reviewer: Bert Brunekreef

Reviewer’s report:

this paper summarizes a number of US studies that have looked to the correlation between personal PM measurements and ambient PM (measured or modeled) from a spatial rather than a temporal point of view.

Personal measurements of PM2.5 and/or sulfate were used. Monthly averages were calculated, and modeled ambient concentrations at the home address were found to be more closely associated with personal measurements than measurements from the nearest monitoring station.

Comments:

Spatial variation in (long term, let's say annual) air pollution is usually rather smaller than temporal variation in the short run, typically daily averages used in time series studies. Validation studies of the spatial, long term exposures used in many epidemiology studies is therefore both needed and challenging. Quite a few studies of the temporal correlations have been published, usually focusing on daily averages as in much of the epidemiology. In contrast, no personal exposure studies have measured for a whole year, so that the validation signal (personal long term PM) itself is subject to error which is usually not well quantified. The authors in this paper have calculated monthly means, and have done some sensitivity work by using the exact dates of the personal measurements only. I think they should do more with the data at hand to quantify the error associated with not having true long term personal PM data. First of all, the monthly means will vary over the year(s) so that part of the reported correlations are temporal not spatial. Can the authors estimate how large the temporal contribution is?

Second, an alternative approach, used in many LUR studies which also suffer from incomplete long term coverage is to calculate unbiased annual means with either the ratio or the difference method, using the signal from one or more continuously operating ambient monitors as input for the calibration. As the authors have the raw data, I think it is both feasible and very useful to do this.

Personal measurements of PM are difficult technically, and (as all measurements) associated with a purely technical measurement error which may be substantial compared to the between-subject variations of the quasi-long term measurement data themselves – especially when the remaining temporal component has been completely removed from the data. Would be useful to have a discussion of this, and an estimate of how important this might be.

Especially in the RIOPA centers, the total personal PM2.5 sd was much larger
than the nearest monitor or model predicted PM2.5 sds. One suspects again that this may be related to insufficient filtering out of the temporal component, as RIOPA had 4 hours of personal measurements only. It could also be related to indoor sources but the tables and figures give us no information on smoking/ETS which should be the biggest contributor in most cases.

Not having data on smoking/ETS in table A4 seems to severely limit the analysis of factors which might explain differences in calibration factors between cities. Another obvious candidate is the ambient and personal distribution sd: one can easily imagine that in populations with little variation in one or more of these distributions, correlations between ambient and personal PM are decreased because of a larger role of measurement error.

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

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

no competing interest