Author's response to reviews

Title: The impact of imprecisely measured covariates on estimating gene-environment interactions

Authors:

Darren C Greenwood (d.c.greenwood@leeds.ac.uk)
Mark S Gilthorpe (m.s.gilthorpe@leeds.ac.uk)
Janet E Cade (j.e.cade@leeds.ac.uk)

Version: 3 Date: 9 February 2006

Author's response to reviews: see over
We have addressed the few remaining comments of the reviewers in the text, responding to them point by point below. We hope that this revised version is satisfactory and we look forward to hearing your decision soon.

Robert Lyles

Major compulsory revisions.

none

Minor essential revisions.

1. What we have referred to as “standard errors” in the text and tables are as the reviewer suggests, the empirical standard deviations of the estimates. We accept the distinction drawn by the reviewer, and have changed the text accordingly to ensure the meaning is clear.

2. As suggested, we have now included the number of simulations and number of observations in each simulated dataset in the footnote of table 1.

3. We have made the suggested change to the wording of the first sentence of paragraph (i) on page 14.

Discretionary revisions.

none
Gita Mishra

Major compulsory revisions.
none

Minor essential revisions.
none

Discretionary revisions.
none
Ian White

Major compulsory revisions.

1)  
   a) The unclear sentence has been removed because we had already made the point more clearly in the preceding text.  
   b) We tried to discuss possible reasons for the findings from the illustrative example, and have kept the explanation (a) that the reviewer liked, whilst clarifying that this is based on a very simple example.  
   c) We have removed the sentence the reviewer disagreed with.  
   d) We have removed the sentence in the discussion that the reviewer disagreed with.

Minor essential revisions.

2) We have now qualified the statements made in the first two paragraphs of the discussion as requested.

3) When a continuous variable measured with error is multiplied by a binary dummy variable to indicate homozygosity, assumed measured without error, and coded 0 and 1 to form the interaction term, the observations multiplied by 0 are fixed 0 with no error attached. Using standard regression calibration treating the interaction term as a new variable irrespective of its meaning, the fitted values are not fixed to be zero, at least not the way I was doing it before. But they should be fixed at zero if you want to interpret things as you expect to. It makes sense that they should be fixed at zero if you consider that using interaction terms is analogous to modelling subgroups. One model will be \( y = a + b_1 x_1 \), the other \( y = A + B_1 x_1 + B_2 x_2 + B_3 x_1 x_2 \). If the fitted values from the regression calibration were not zero, then an element of additional measurement error would be introduced to the first equation, when in fact it should all be in the \( x_1 \). I’ve not said this very well, but it seems the appropriate way to handle it to me, in the same way as true vegetarians do not eat meat, so their meat intake should be taken to be zero, irrespective of any measurement error in the meat intake of meat-eaters. From another perspective, my understanding is that the sort of skewness introduced by this type of interaction term violates the principals of regression calibration (see comments in Robert Lyles’ review).

4)  
   a) We accept that this sentence is awkward, and that it is important to state that is the mean, but we wanted to draw distinction between what we’ve termed the “coefficient estimate” and the “ratio estimate”. That’s because our findings appear to contradict those of Wong & Day because they express theirs in terms of ratios. We have inserted “mean” in the sentences to clarify, and hope that this satisfies your point.  
   b) These lines have been simplified.

5)  
   a) Done.  
   b) Done.
c) Done.

d) Done.

e) Done.

Discretionary revisions.

none