Author's response to reviews

Title: Environmental exposures and their genetic or environmental contribution to depression and fatigue: a twin study in Sri Lanka

Authors:

Harriet A Ball (harriet.ball@kcl.ac.uk)
Sisira H Siribaddana (nipuna@stmail.lk)
Athula Sumathipala (thula.sumathipala@kcl.ac.uk)
Yulia Kovas (yulia.kovas@kcl.ac.uk)
Nick Glozier (nglozier@usyd.edu.au)
Peter McGuffin (peter.mcguffin@kcl.ac.uk)
Matthew Hotopf (matthew.hotopf@kcl.ac.uk)

Version: 2 **Date:** 28 October 2009

Author's response to reviews: see over

The Editor
BMC Psychiatry

29th October 2009

Dear BMC Psychiatry Editor,

Please find attached the revised version of our article 'Environmental exposures and their genetic or environmental contribution to depression and fatigue: a twin study in Sri Lanka' for your consideration for publication in *BMC Psychiatry*. We have addressed all of the points raised by the reviewers in turn.

Reviewer 1: Margaret Gatz

This reviewer was concerned that the problem of reverse causation was not emphasised enough, both in our reference to 'depression' rather than 'history of depression' and the extent of the implications of adjusting for past-year work. We have made the following changes:

- We have acknowledged that adjusting for work does not fully account for the reverse causation problem, by adding the following sentence on p15: "This finding reduces the likelihood of one pathway of reverse causation, but there still could be other reverse causation pathways, or the effect on work could have been longer ago than the previous year."
- We acknowledge that it is important for the reader to understand that depression is a lifetime-ever assessment, i.e. of their history of depression rather than current depression. We have clarified the situation through the following:
    - In the Methods-Data Collection section (p6), we have added the following sentence: "The current analyses pertain to lifetime-ever history of depression, rather than current depression."
    - We have edited a further 26 sentences, so that they refer to history of depression rather than just depression. This means the reader is made aware of the issue every time it comes up. Including "A history of depression was present in 11.1% of the sample…" on p13.

Reviewer 1 also pointed out that we had not described the age structure of the sample. The following sentence describing demographic characteristics has been added on p5: "Among men, the mean age was 33 years (s.d. 13); among women the mean age was 35 years (s.d. 14); 46% of the participants were men."

Reviewer 2: Marieke Wichers

1. This reviewer wondered why we had not used a continuous measure of depression, for example the symptom count. Unfortunately, whilst DSM-IV clinical symptoms of depression would provide a continuous measure, this would be skewed (many participants scored zero). The structural equation models require continuous variables to be normally-distributed, which is why we chose the categorical variable, and ran models appropriate to this data.

2. This reviewer also queried whether we could have used GE interaction models. Unfortunately, because we are tied to using categorical variables we do not have enough power to run such models.

Reviewer 3: Yoon-Mi M Hur

This reviewer questioned why we had used logistic regression models, rather than structural equation models (SEM), to examine the ACE influences on the overlap between exposure and depression or fatigue. We appreciate this concern, and the fact that SEM allows percentage effect sizes to be calculated. However, we think the logistic regression models are preferable in this instance. We have already described our reasons on p12 of the manuscript, but briefly they are:

- Under certain circumstances, MZ differences approaches can have greater power than a full SEM model (see reference 29 of the manuscript), and we found we were able to run our analyses on a narrower definition of depression than was possible using SEM. This meant we could potentially pick up more severe cases that might have a stronger association with important causal risk factors.
- The logistic regression approach is more accessible to an audience likely to be familiar with regression models of some sort but not structural equation models. It also allows a very intuitive method of adding in potential confounding factors.
- The logistic regression approach requires no assumption of bivariate normality (and in our epidemiological analyses of the same sample – reference 20 of the manuscript – the relationship between the exposures and depression was stepped rather than smooth across the distribution of exposures). This wouldn't preclude running SEM on these data, but it would make their interpretation more difficult.

Finally, although SEM would provide an estimate of effect size, our experience of running similar categorical models on samples of this size suggests the confidence intervals would be quite wide. This means the useful information gained would not be far different from simply finding the presence or absence of an effect.[hab1]

This reviewer also requested that we use SEM models to look for gene-environment correlation and gene-environment interaction. However, she notes that our models already test GE correlation (e.g. the data suggest GE correlation can partially explain the association between depression and life events in women). Unfortunately, due to the categorical nature of the data, which results in lower power to test complex models, we would not be able to run GE interaction models. We have now acknowledged that our analyses do not consider GE interaction in the Limitations section (p21): "Finally, although our analyses examining the overlap between measured exposures and fatigue or depression outcome looked for correlations between genes and environments, our assessment of the heritability of depression and fatigue did not assess potential interactions between genes and environments."

This reviewer also pointed out the lack of information about the age distribution, which we have included in response to reviewer 1's comment above.

Finally, this reviewer pointed out we needed to better conceptualise the correlations between the 4 measured environments before moving on to analysing their relationship

with depression and fatigue. These correlations have now been reported in a footnote to Table 1, and summarised in the text on p13 (they ranged from 0.38 to -0.12). These correlations mean that associations between depression/fatigue and more than one exposure could be driven by just one of the exposures. For this reason, we already took care to include the following in the original version of the manuscript:

- The phenotypic (within person) associations of each exposure with depression/fatigue, as reported in Table 1, are independent of the other 3 exposures.
- Because three exposures contributed to nonshared environmental mediation on history of depression in men, we describe an additional model examining these 'E' effects independently of one another (p15).

The manuscript is intended as a full-length article and is not being submitted for publication elsewhere. All authors have read and approved the current version of the manuscript. Contact details for the corresponding author are as follows:

Harriet Ball, SGDP Centre, Box P080, Institute of Psychiatry, De Crespigny Park, London SE5 8AF. Tel: (0)207 848 5415, fax: (0)207 848 0866, email: harriet.ball@kcl.ac.uk.

We look forward to hearing from you.

Yours sincerely,

Harriet A. Ball, Sisira Siribaddana, Athula Sumathipala, Yulia Kovas, Nick Glozier, Peter McGuffin & Matthew Hotopf