Author's response to reviews

Title: Antecedents of teenage pregnancy from a 14-year follow-up study using data linkage

Authors:

Jennifer Gaudie (shenny@iinet.net.au)
Francis Mitrou (francism@ichr.uwa.edu.au)
David Lawrence (D.Lawrence@curtin.edu.au)
Fiona J Stanley (fiona@ichr.uwa.edu.au)
Sven R Silburn (sven@ichr.uwa.edu.au)
Stephen R Zubrick (steve@ichr.uwa.edu.au)

Version: 2 Date: 23 December 2009

Author's response to reviews: see over
We would like to thank the three reviewers for their helpful comments on the paper. We have revised the manuscript in line with the reviewers’ suggestions. We feel that the paper has been significantly improved thanks to the valuable feedback received from each of the reviewers. We have detailed below the specific changes that have been made to the manuscript in response to each of the individual points raised by the reviewers.

**Reviewer 1**

*Clearly state the study objectives in the end of background section or beginning of methods section.*

We have clarified the study objectives and stated these at the end of the background section.

*Clarify what is meant by clustered representative probability sample either by explaining or providing a reference for this classification.*

We have added in further explanation of the nature of the sample design, along with a citation.

*Explain all statistical analysis conducted (e.g., bivariate analysis)*

We have expanded the description of the analysis methods, and we have removed the confusing term “bivariate analysis” from the discussion.

*Present values for statistical tests and associated p-values for the statistical analysis conducted.*

We have added in the test statistic and p-values for tests of difference in the hazard curves for girls with both aggressive and delinquent behaviours, girls with either behaviour, and girls with neither.

*Please situate Australia in OECD countries in terms of teenage pregnancies. Some figures on current numbers and trends would be advisable.*

Australia’s ranking among OECD countries has been included in the introduction section.

*Write the meaning of WACHS.*

The acronym WACHS (Western Australian Child Health Survey) is now defined both in the abstract and the background section.

*Include % along with counts*

All percentages shown in the results section and in the table are weighted percentages, estimated accounting for the complex sample design. We have provided raw counts of numbers of outcomes to show the size of the study, but have not provided unweighted percentages to avoid confusion between unweighted and weighted estimates of the same quantity. For instance, we observed 155 teenage pregnancies in 1,374 girls, which is 11.3% unweighted. However, after accounting for the sample design we estimate the proportion of WA girls aged 4-16 in 1993 who went on to have a teenage pregnancy as 10.5%. The weighted estimate is the preferred estimate of proportion of teenage pregnancies.

*Use another notation for confidence intervals.*

While there does not appear to be a standard for the presentation of confidence intervals in BMC journals we have chosen the style that appears to be used most frequently in recent BMC Public
Health papers. We would be happy to take advice from the editors if we have misunderstood the preferred style for the journal.

Reviewer 2

Please add a summary table of the sample demographics from the original survey

We have added an additional table as Table 1, showing the demographic characteristics of the girls from the survey.

Was participant consent required to access records via the WADLU?

Participant consent was not obtained, and the data linkage was undertaken under established protocols for data linkage where consent is not provided. These protocols call for a separation of identifying information from research data, and the creation and analysis of de-identified files. As 15 years had passed since the 1993 survey it was not possible to track down the study participants and obtain their consent for the study. We have added an additional paragraph into the methods section to describe the consent issues and the protocols for ensuring participant confidentiality.

Given that the total sample size at 14 year follow-up is 155, please indicate whether this is a large enough sample to determine statistical significance

The total sample size for the study was 1,374 girls. Of these girls 155 had a pregnancy (either birth or abortion) prior to age 20 which was the main outcome measure for the study. We regret the confusion on this issue, and we have amended the description of the study to clarify that all 1,374 girls were followed-up via data linkage.

We have added a note on statistical power to the methods section and to the limitations section of the discussion.

‘a de-dentified confidentialised file’ - what does this mean?

We have added a paragraph in the methods section describing in detail the creation of the de-identified file. As suggested by the reviewer we have substituted the word ‘anonymised’ for ‘confidentialised’.

P11 par 3 “The proportion of girls who became teenage mothers ...” could this be rephrased “the proportion of girls who became pregnant and continued to parenthood...”

This sentence has been rephrased as suggested.

P12 par 2 I am uneasy about the use of the word “if”

This sentence has been rephrased as suggested.

P12 par 4 I don’t like the use of the phrase “the girl”. Can this be rephrased?

This paragraph has been rewritten, and “the girl” has been removed.

Can the authors rephrase the research questions to give a sense of direction
We have rephrased the statement of the objectives of the study in the last paragraph of the background section to reflect the three aspects of the study highlighted by the reviewer. We agree that this is a better way to describe the aims of the paper.

*Does the data allow the researchers to identify demographics of the sample at 14 year follow-up?*

Unfortunately no. The demographic and other predictor variables collected during the 1993 survey cannot be updated from the information on the linked data set, and we have no way of assessing social mobility in the sample since 1993.

*P18 par 1 this is an awkward phrase.*

The sentence has been rewritten in positive terms and the double negative has been removed.

**Reviewer 3**

*Was there any need to take into account the household-based nature of WACHS when extrapolating your findings to the general population?*

Yes, the complex nature of the survey design was taken into account in both the weighted estimates shown in Table 1 and in the proportional hazards regression model. We have expanded the description of the analysis methods to more clearly spell out how the survey design was accounted for in the analysis.

*How did you evaluate power for multivariable analysis? Is it possible that the lack of contribution of some variables to the outcome is due to low power?*

We have added a section on power into the methods section. We estimated the study would have about 80% power to detect a hazard ratio of 1.7 or greater for a risk factor with prevalence of 10 percent. The study would have lower power to detect associations for rarely occurring risk factors or risk factors associated with small effect sizes. We have noted this limitation in the discussion.

*Having an idea of the spline functions fitted to your time-to-event data may be useful for researchers trying to reproduce your results.*

Our approach to fitting a spline to the relationship with age used the generalised additive models framework. This is a non-parametric method, and the resulting curve is not expressed in a functional form that can be written down in the form of an equation. The standard approach to describing the shape of a curve fitted using GAMs is to graph the results, which we have done in figures 1 and 2.