Reviewer's report

**Title:** Health service use in families where children enter public care: a nested case control study using the General Practice Research Database

**Version:** 1  **Date:** 23 November 2011

**Reviewer:** Shirley Porterfield

**Reviewer's report:**

This article examines factors associated with the entrance of children into public care in England. The General Practice Research Database is used to identify mother-child dyads where the child was placed in public care. These dyads are matched on the basis of child age and sex with four other mother-child dyads from the same practice.

Overall, this article feels like a fishing expedition. The reader is not quite sure what the authors already know on this topic, nor what they think they might find from this analysis. The research is characterized as exploratory, but this is a topic about which much is already known. It would be helpful if the authors would identify what is already known, what is not known, what this paper will examine (specifically and in detail), and then focus on the results that are new, rather than giving equal focus to known and new results. The pieces are mostly all here, but they're not arranged as well as they could be. I consider this rearrangement and a revision of the Background section (detailed in the next paragraph) to be Major Compulsory Revisions.

**Major Compulsory Revisions:**

1. The Background section of this paper is very terse and needs to be fleshed out, with more reference to existing literature on this topic, in order to motivate the study. For example, the first sentence mentions significant health problems of children at the point of their entry into public care. What kind of health problems are these? Are they physical health problems? Mental health problems? The second paragraph suggests that several studies on this topic have been done, though few in the UK. What did these other studies find? What do the authors of this study expect to find? There is a statement of purpose, but it’s very general, and is not followed by any hypotheses or expectations regarding potential findings. Interestingly many of the earlier findings, which I would expect to see in a background section, show up instead in the discussion section at the end. Move those to the front so that you’ve fully outlined the issue you plan to examine and made clear what will be added by this study.

**Minor Essential Revisions:**

2. I am not familiar with this data set, but it appears to be appropriate to this topic. It would be helpful to the reader if all information pertaining to the GPRD dataset were in the same place and in the methods section. Currently there are
three different sections with some variation of the heading “The GPRD dataset.” Put these altogether.

3. The most surprising thing about the dataset is the small sample size eventually arrived at by the authors. It’s unclear from the figure how the dataset went from 2,954 case dyads to 147 case dyads. The figure only shows the progression from 370 case dyads to 147. What happened to the rest?

4. In the section “identifying variables for analysis”, the authors mention several variables that have been found in other research to be associated with the entrance of children into public care. Those variables should be mentioned (with citation), but should not be the focus of this paper unless a case is made (in the background section) that the effect of these variables on entrance into public care in the UK is not known AND there is some reasonable argument made that unique circumstances in the UK suggest they might have different effects here. There are several variables that may be important that are left out, I’m assuming because the dataset doesn’t include these measures. Variables such as household or family structure (number of adults in the household), disability in the family, and/or age of the child at entry into care are not included (though age of child at entry into care is included in figure 2). Are the clinical event codes referred to something like ICD-9 or 10 codes, or are they unique to the NHS?

5. The discussion of the statistical analysis in the methods section mixes the two models estimated. It would help the reader if these two analyses were described separately. The terminology here is somewhat confusing as well, particularly the “individual variables conditional logistic regression.” The “forward selection” model is a form of stepwise regression, a term which might help some readers to better follow the method of analysis. The methods are appropriate to the question being addressed.

6. The Discussion section should highlight unique and/or important findings, such as maternal mental illness, possibly SES (though this relationship appears well known from previous literature) as the discussion in this section is interesting, other maternal factors, and child related factors. Move the discussion of variables that were not identified as risk factors into the methods section as part of the discussion of the dataset.

Discretionary Revisions:

7. The methods appear to be sound, but not as well described as I would like. Also, some of the terminology may be confusing to those not from the UK. Some of the decisions about case and control definitions are not well explained. For example, why look at health services use over a period of 12 months? Why choose 4 control cases and why only match them by the age and sex of the child? The one variable mentioned as important in the introduction was socio-economic status (SES) so why not also use SES to match? Clearly a case can be made for each of these decisions, but I don’t see it until I get to the discussion section. Reasons for these decisions should be discussed in the methods section.
8. The authors mention that not many fathers are registered at the same practice as their children, while there is the suggestion that most mothers are. What percent of mothers are registered at the same practice as their children? More than half? Ninety-five percent? This would make a difference in generalizability of results. It’s interesting that there doesn’t appear to be a pointer from child’s record to the records of mom and dad, though clearly this is not something the authors can control.

9. The small section (actually 2 small sections) on data validation is interesting, but more appropriately belongs in a footnote or endnote.

10. In the Conclusion, remind the reader what is unique about the results from this article (as well as the novel approach).

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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

Declaration of competing interests:

I declare that I have no competing interests