Reviewer's report

Title: Adjusting for under-identification of Aboriginal and Torres Strait Islander births in time series produced from birth records: using record linkage of survey data and administrative data sources

Version: 1 Date: 26 April 2012

Reviewer: Sanja Lujic

Reviewer’s report:

Thanks to the authors for their submission outlining the under-identification of Aboriginal and Torres Strait Islander births in Western Australian administrative databases. Aboriginal under enumeration is highlighted in several governmental reports and journal articles, and is an area of importance. Notwithstanding it’s topical nature, there is a lack of validations studies that aim to evaluate Aboriginal recording in administrative datasets. The authors have shown that Aboriginal birth records in Western Australia are under identified, as well as providing useful information about biases introduced when examining low birth weight and gestational age information.

My comments are structured as per the Journals format, for clarity purposes. Points of clarification are raised in the formatted section (1-9) as well as the major/minor/discretionary revision section at the end.

1. Is the question posed by the authors well defined?

Yes, the authors have clearly described the main objective of the study.

2. Are the methods appropriate and well described?

Yes, the methods are appropriate. However I would have liked to see more information regarding:

a) Phrasing of the questions about Indigenous status in the Register of Births, Midwives’ Notification System and WAACHS survey.

b) A reference and 1-2 sentence description about the level of relative isolation, socioeconomic disadvantage. Has the latter derived from the ABS SEIFA?

c) I also suggest reporting sensitivity, specificity and agreement measures (please see Major compulsory revisions section below).

3. Are the data sound?

Yes, the data are representative of the state-wide population of children aged 0 to 17 years.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

I think this is fine. Some comments for improvement are outlined in revision sections below.
5. Are the discussion and conclusions well balanced and adequately supported by the data?
I feel that the discussion section could benefit from a slight rewrite, as well as comparing the study results to the Australian literature – please see a list of relevant papers in the revision section below.

6. Are limitations of the work clearly stated?
Yes. Perhaps a mention about non-response to WAACHS survey could be beneficial. Are the non-responders more likely to be families with poorer child outcomes, living in more disadvantaged areas, and living in more remote communities?

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
The paper would benefit from the inclusion of recent relevant articles (please see Major compulsory revision for details). It might also be worthwhile indicated how the Midwives’ data is used currently in WA, in terms of the standard reporting (e.g. Mothers and babies reports in New South Wales). Does WA Health publish any regular reports?

As a means of context, what % of populations/children in WA are Indigenous? And where do the majority of Indigenous people live? Have there been any other studies in WA that looked at Indigenous reporting?

8. Do the title and abstract accurately convey what has been found?
Yes the title clearly reflects the topic under investigation.

9. Is the writing acceptable?
The article is well written, however some editing is required. Please see the major compulsory revisions section below.

Major Compulsory Revisions

a) Please report sensitivities, specificities, positive predictive values and negative predictive values. The tables with weighted crude numbers are informative but it might be worthwhile incorporating other measures.

b) The paper needs some editing, to remove excessive (redundant) detail and improve continuity. Although the authors have given a comprehensive outline of the methods and findings, perhaps shortening certain sections might improve readability and clarity of the paper. This could be achieved by removing superfluous and repetitive sentences, usually found in both the methods and results sections. Some examples are outlined in the Minor Essential Revisions section.

c) It could be useful to add information about other validation studies in Australia or overseas in the Background/Introduction section, especially relating to under-identification of Aboriginal status in administrative datasets. At present, a literature review is missing in this section. The authors might like to consult some
of the references below:

• Xu F, Sullivan EA, Madden RC, Black D, Jackson Pulver LR. Improvement of maternal Aboriginality in NSW birth data. BMC Medical Research Methodology 2012; 12:8


• ABS 2006. Births, Australia, 2006. ABS cat. no. 3301.0. Canberra: ABS. (95% of Indigenous mothers were identified as Indigenous in the WA Registry of Births from 2002 to 2006)


• Centre for Epidemiology and Research: New South Wales Mothers and Babies 2006. NSW Public Health Bull 2009, 20(s-1):1-156


Some other references relating to the validation of maternal morbidity are:


• Roberts CL, Bell JC, Ford JB, Morris JM. Monitoring the quality of maternity care – how well are labour and delivery events reported in population health data? Paed Perinatal Epidemiol 2009; 23:144-152

d) It is not clear how information from siblings is treated in the analysis. Are the models adjusted for obvious clustering effects (children from the same mother)?
Minor Essential Revisions

From the Methods section:

a) What number and proportion of 5,289 children had mothers as carers?
b) How is level of relative isolation measured?
c) What does family structure comprise of?
d) How is relative socioeconomic disadvantage measured? Which geographical level was used to attach SEIFA scores (CD, SLA, LGA, postcode)? Part of this information is included on page 20 in the Results section; I suggest moving this information to the Methods section.
e) It is unclear which data source(s) was used to obtain birth weight, gestational age, and rented accommodation and test scores. Presumably the first two are obtained from the Midwives notification system, and the latter two from WAACHS? Perhaps stating this in the methods could be useful for the reader.
f) Adjusting administrative-based time series section (page 13) could potentially be broken into 2 sections, as it mentions different models/results using different data sources:
   a. Factors contributing to correct recording
   b. Adjustments of time series data
g) Some of the information in the background section of the paper might be better placed in the Methods section. For example, paragraph 2 on page 7 can be moved to Data Sources section in the Methods.
h) Which statistical package was used to analyse the data?
i) Perhaps a diagrammatic representation of data sources and overlaps can be incorporated to distinguish which data sources/years/available information and respondent numbers were available for analysis at different stages.

From the Results section:

a) Page 15 – first sentence is repetitive, please remove. Potentially remove sentences 2, 3 and 4 and combine that information with sentence 5.
b) Page 16, first paragraph – this information is better suited in the Discussion section. First sentence of the second paragraph is repetitive, please remove. Last sentence – father’s status is missing in 33% of records, not 31%.
c) What % of children had either Indigenous mother or father using Register of Births?
d) Page 18, inconsistent identification section – first 2-3 sentences are repetitive, please remove.
e) Page 20 – first paragraph should be moved to the Methods section. Second sentence can be deleted, as it is repetitive.
f) Please fix figures – they are hard to read at the present. Perhaps use black/white/grey instead of colour, and solid and dashed lines? What do the lines
represent (e.g. Lowess smooth)?

From the Discussion section:

a) The authors might consider structuring the discussion section as follows:
1. Re-iterate the main findings of the study
2. Discuss the findings in light to previous research
3. Outline strengths and weaknesses of the study
4. Potential future work

Discretionary Revisions

a) Ethics approval information generally appears in the last section of Methods
b) Perhaps Aboriginal and/of Torres Strait Islander can be mentioned first up, and for the remainder of the paper, this could be replaced by Aboriginal for the ease of the readership? Perhaps the same could be applied to terms ‘WA Register of Births’ and ‘WA Midwives’ Notification System’, if possible?

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests