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: 16 January 2012

Reviewer: Lee Taylor

Reviewer’s report:

Major compulsory Revisions

1. The use of the terms “Indigenous” or “Aboriginal” or “Aboriginal and/or Torres Strait Islander” should be made consistent throughout the paper.

2. Methods. A definition of an Aboriginal and/or Torres Strait Islander infant needs to be stated. From the remainder of the paper it appears that two concepts are used: infant reported to be born to an Aboriginal and/or Torres Strait Islander mother or father; and an infant reported to be born to an Aboriginal and/or Torres Strait Islander mother. Which analyses were carried out in relation to each concept needs to be clearly described.

3. Results. Comparison of Indigenous status between the WAACHS and the WA Register of Births and the WA Midwives Notification System.

   a) Paragraph 1. Much of the first part of this paragraph refers to methods that would sit better in the Methods section of the paper. I would suggest the text up to the sentence reporting the first results shown in Table 1 be moved to the Methods. The part of paragraph commencing “As the identifications in the survey were all made…” to the end of the paragraph includes discussion rather than results. Discussion should be moved to the Discussion section of the paper.

   b) Paragraph 2. The first sentence repeats information in the Methods and should be removed. The second and first part of the third sentence should be moved to the Methods and only the results reported in this section.

   c) Paragraph 3. This paragraph does not refer to any results and should be removed.

4. Results. Characteristics of children who would be identified differently between administrative and survey data sources.

   a) Paragraph 1. Sentence 1 and 2 refer to “the survey”—which of the two survey sources is this? Sentence 2 also refers to “both” sources. However, there are three data sources mentioned in this paragraph— it is a bit confusing where “administrative data sources” are referred to as a single “data source”. The wording here should be clarified.

   b) Paragraph 2. References to “the survey” and “both” as per comment 5. above.

5. Results. Inconsistent identification of non-Aboriginal children in the 1993 WA Child Health Survey. The first part of this paragraph (up to “as shown in Table 5”)
repeats information that was previously presented in the Methods. This should be replaced by a short introductory sentence or phrase.

6. Results. Time series of low birth weight and low gestational age babies. Information on methods that has previously been described in the Methods section should be removed.

7. Discussion. As per 3.a) above, it is a bit confusing where “administrative data sources” are referred to as a single “data source”. The wording here should be clarified.

8. Discussion. Paragraph 10. Last sentence It is not clear what is meant by “neonatal status” as an outcome variable. Does this refer to the low birth weight and prematurity outcomes previously discussed? If so, then either a definition of neonatal status should be included in the Methods or the outcomes should be specifically stated here.

Minor Essential Revisions

1. On occasion, the term “demographic data” is used to refer to “reporting of Aboriginal and Torres Strait Islander peoples” (e.g. Background, paragraph 4). The term “demographic” should be replaced by “Aboriginal and Torres Strait Islander peoples” or a similar descriptive term where relevant.

2. Background, paragraph 4. In order to clarify the study aims, the term “administrative data collection/s” should be replaced or supplemented by the name/s of the data collection/s.

3. Background, paragraph 8. “As there is…”, should be replaced by “As there are…”.

4. Methods, Analysis methods, paragraph 1. The last sentence refers to “children and young people”. However, this study concerns neonates. The wording here needs clarification.

5. Methods, Analysis Methods, paragraph 3. Information on the measurement of “level of isolation of residential address” and “family structure” should be included in the Methods.

6. Methods, Adjusting administrative-based time series, last paragraph. “…summing the probabilities within categories of neonatal outcomes” should be changed to “…summing the relevant probabilities” as the outcomes have been previously referred to in the same sentence.

7. Results, Characteristics of children who would be identified differently between administrative and survey data sources, paragraph 1. The first sentence refers to “outcomes”. However, the variables listed are not all outcomes. The term “outcomes” should be changed to “indicators” or “measures” or a similar term.

8. Discussion. The heading “Limitations” should be removed as there are no other headings in the Discussion. An introductory phrase could be included at the beginning of the next paragraph.

9. Tables and Figures. The location and relevant years should be included in the title. For example, Table 1 should be modified to read “….weighted survey
estimates, Western Australia [YEAR] to [YEAR].

10. Tables and Figures. Footnotes to each table indicating the scope of data reported in the table would be very helpful e.g. Linked records of [X] and [Y] datasets, or All records of [X] dataset and linked records of [Y] dataset. For tables where weighted survey data are reported, a footnote should refer to the number of records in the survey that are represented in the table.

11. Tables 1, 2 and 5. The inclusion of row or column percentages as relevant will facilitate interpretation of the data.

12. Tables 3 and 4. The numbers in these tables are presumably all percentages. This needs to be clarified in the column headings or elsewhere. It needs to be clear whether the percentages are row or column percentages or if there is some other denominator.

13. Table 4. The table would be more readable if the confidence intervals were located in columns alongside the point estimates, rather than below, with a relevant column heading.

Discretionary Revisions

1. Methods, Data Sources, paragraph 4. This paragraph may be deleted and the references moved to the previous paragraph, as this single-sentence paragraph adds little to the paper.

2. Discussion, paragraph 3, last sentence. The word “less” could be replaced with “fewer”.

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 have no competing interests.