**Author's response to reviews**

**Title:** Record linkage to obtain birth outcomes for the evaluation of screening biomarkers in pregnancy: a feasibility study

**Authors:**

Samantha J Lain (samlain@med.usyd.edu.au)
Charles S Algert (calgert@med.usyd.edu.au)
Vitomir Tasevski (vtasevsk@nsccahs.health.nsw.gov.au)
Jonathan M Morris (jmorris@med.usyd.edu.au)
Christine L Roberts (clroberts@med.usyd.edu.au)

**Version:** 2  **Date:** 11 May 2009

**Author's response to reviews:** see over
Re MS: 1589263631243316
Revised title  Record linkage to obtain birth outcomes for the evaluation of screening biomarkers in pregnancy: a feasibility study
Authors: Samantha J Lain, Charles S Algert, Vitomir Tasevski, Jonathan M Morris and Christine L Roberts

Thank you for the opportunity to respond to the reviewer’s comments on the above manuscript. We have been delayed in our response to the reviewer comments, as the author originally designated as the corresponding author has been away on maternity leave. We apologise for not responding sooner. Please extend our thanks to the reviewer for their interest and comments. Changes to the original manuscript, where discussed below, have been made using Track Changes.

We have discussed below, have been made using Track Changes

Methods
“It is not clear who and how de-identified data set of the laboratory results to be able to match with birth outcomes”
The methods have been revised to clarify the separation of identified data (for linkage) and de-identified data for research.

Fetal loss and stillbirth
In response, we have added “stillbirth” as one of the stated outcomes in the Methods section of the Abstract (last sentence of paragraph), so that it is consistent with the presentation in the main Methods section and in the Results (Table 2). As the reviewer points out, the usage of the word stillbirth varies. Researchers have used 22, 23 and 24 weeks as a cutoff to make a distinction between fetal loss and stillbirth. During the period of our study, 24 weeks gestation was considered as the edge of viability so we have used this as a threshold.

It should be stated how the data ... are audited (ref 2, 29)
The referenced published articles explain how the datasets were audited. However we have now indicated in the discussion that the PHDS have been validated in studies that compared PHDS reports to the corresponding medical records.

“Where (sic) the used data cleaned?”
The population data are cleaned, logic checked and duplicates are removed by the Department of Health before release, and the PaLMS laboratory data had been checked by those responsible for the Down screening program.

Preterm birth – 78 births in Table 2 and 70 in text
The text (page 6) states “Among the livebirths ...”, before detailing these 70 births. Table 2 shows all preterm births (live and stillbirths). In response to the reviewer, we have added a footnote to Table 2 clarifying that the preterm birth numbers include preterm stillbirths.

“Is it true that only 75.8% of preterm birth reports are correct?”
“Is it true that only 98.3% of stillbirth reports are correct?”
As the text states, 75.8% is the sensitivity of reporting of preterm birth categories in the hospital data. Sensitivity refers to the percentage of true cases identified in a report. The percentage of reports that are correct is more akin to the positive predictive value, or the agreement proportion. For this study, both hospital records
with preterm reports and gestational age reported on the Midwives Data Collection were used, so that a high percentage of cases could be identified. In response to the reviewer's comment, we have edited this paragraph (final paragraph of the Discussion) to simplify and make clear that reports from both datasets can be used to improve ascertainment of outcomes.

What is probabilistic linkage? What is translational research?
We have added a brief statement to the methods that describes probabilistic linkage and added references to a comprehensive reviews on this complex topic. The conclusion paragraph has been revised to make it clear that we were referring to translating research findings into clinical practice (translational research).

Discussion
“Traditional methods will be needed ...”
The reviewer makes the point that not all countries have extensive population health databases which can be or have been linked. Traditional methods of patient follow-up and medical review would then be the only option.
In response to the reviewer’s comments, we have
- qualified the Discussion by adding the word “existing” to mention of population health datasets in the first and fourth paragraphs
- added to the limitations that using this methodology requires access to routinely collected population health data
- indicated that the comparative studies are all prospective, 2nd paragraph of the discussion
- added a descriptive of a traditional prospective approach to this type of research
(p11)

“who performed the blood sample analyses?”
The PaLMs laboratory is a business unit of the public hospital network in northern Sydney. We have added wording to the second sentence of the Methods, to clarify this.

SUGGESTED REVISIONS
“possibility of change in the title”
The idea of using linked data can be applied generically, but we take the reviewer’s point about this being specifically about lab (pathology) data. In response we have changed the title to:
“Record linkage to obtain birth outcomes for the evaluation of screening biomarkers in pregnancy: a feasibility study”
“It is not clear who and how de-identified data set of the laboratory results to be able to match with birth outcomes Is this really a new method?
Explain what is probabilistic linkage? What is translational research?”
See our previous response to these questions about probabilistic linkage and translational research.

“Who performed laboratory work? acknowledgements?”
We state that PaLMs (the Pacific Laboratory Medicine Services) does the serum analyses, in the second sentence of the Methods.

MINOR ESSENTIAL REVISIONS
“Whole text: explain what you mean by effective, efficacious, efficient, cost efficient, cost effective and use words consistently. Predictive tests; use predictive value of tests. Sample (not spectrum). Screening factors? Screening tests.”
We feel that the normal usage of these words applies here, and that an expanded discussion of their meanings would not be useful. This paper is not meant to present a detailed economic analysis, the focus is on the reliability of the scientific results. As to the other comments here, they are somewhat unclear, since they lack specific references to the text.

“Abstract: the aim of the study is to test feasibility. The results should be written about that.”
Yes, the Results and Discussion are about the feasibility of using the linked data. Feasible in the sense that the results compare well with other studies which have used case review, so the linked data can be reliably used.

“Write about 4 outcomes (fetal loss and stillbirths separately)”
As above, we have added stillbirth as a fourth outcome to the Abstract. In the Methods and Table 2 of the Results, stillbirth (≥24 weeks) was already presented as a separate outcome.

Background: 2nd paragraph: it should be clear that the data cited are not from Australia (ref 5).
In response, we have reworded the last sentence of the paragraph to make this clear.

“It should be stated how the data from the laboratory and the data in health databases are audited.”
See the response to the query about data audits, on page 1 above.

Methods, 2nd paragraph: pregnant women were expected to deliver (not pregnancies); the same in Results.
In response, we have changed “pregnancies” to “women”, in both places.

Table 1 is not appropriately explained in the text (only 0.26, and 0.28 are explained; 0.49 and 0.92 are not). Why?
Our impression is that journals normally frown upon tedious repetition in the main text of data already shown in tables. In response to the author, we have shortened this sentence and used only one of the statistical comparisons as an example.

Table 1: free beta HCG ? differently written in the same table
In response, we have edited the table to make these consistent 5.3% and 3.5%: is it correct that this is not significant difference?
Yes, it is correct. The differences in these proportions (PAPP-A ≤5th percentile threshold) are not statistically significant.

Table 2: 78 preterm births? In the text: 2 + 16 +50 = 70
As discussed in the earlier response (mid-page 1 above) there were 70 preterm livebirths and 8 preterm stillbirths.

The meaning is not clear, please rewrite:
page 9: ?potential pregnancy screening factors
In response, we have added an example of such potential screening factors.
page 9: the sentence that starts ?With a sample size?? .
In response, we have rewritten this sentence.

How did you prove it is less expensive? Was an economic analysis performed? Did you consider expenses for establishing databases, centre for linkage, data protection, auditing all these constantly?
As discussed in our previous responses, this method is useful for existing databases (we added that qualification to the revised text, in response to the reviewer’s comments). We have also added that this methodology requires access to routinely collected population health data. Finally we have added a cost comparison to a contemporary cohort study of pregnancy biomarkers which to date has been allocated funding of AUD$11.4 million

References: cite in the same way
In response, we have made the citation of journal titles consistent in the references

MAJOR COMPULSORY REVISIONS
Define the aim of the study in the Abstract as it is in Background (p5) and change the title (this is not evaluation of screening test but about feasibility).
In response, we have made the study aims in the Abstract identical to the aims of the study in the Background (p5). However, we do not believe that the title needs to be changed any further. The fact that the title ends with “: a feasibility study” should be sufficient to alert readers to the fact that this is primarily a feasibility study of using record linkage to obtain study outcomes.

Discussion: discuss mainly about feasibility; compare results with results of studies done in different ways (do not discuss the predictive value of screening tests).
“Feasible” in this context means that a novel research method can produce reliable results; in this instance, that the predictive values obtained are comparable to other methods. Table 3 is precisely about comparing our results with the results of other traditional cohort studies. The most useful comparison is of effect measures that are independent of the incidence of adverse outcomes in the different study populations, which means a relative measure of effect such as the odds ratio or risk ratio. Threshold sensitivities and PPV’s are also compared in the Discussion since this provides further assurance that our results were in line with other studies, and the data may be useful to other investigators.
2nd paragraph: screening tests are NOT diagnostic tests.
In response, we have changed the word “diagnostic” to “predictive”.

“Is it true that only 75.8% of preterm birth reports are correct?”
“Is it true that only 98.3% of stillbirth reports are correct?”
See page 1 or these responses, where we have already responded to these queries.

Discuss why traditional research methods will still be necessary.
See our earlier response and revisions regarding traditional methods