Reviewer's report

Title: Individual Patient Data meta-analysis of diagnostic and prognostic studies in obstetrics, gynaecology and reproductive medicine.

Version: 2 Date: 21 January 2009

Reviewer: Lelia Duley

Reviewer's report:

The additional detail on methods substantially improves the protocol. However, I have the following comments:

Under methods, the heading ‘Identification and selection of studies’ should remain, as the search strategy is critical. The references cited here as having identified the primary studies were all published in 2002. These searches are therefore not up to date, and a strategy for updating them should be outlined here.

This protocol should also include a list of the studies identified to date for each of the topics. This would give an estimate of the number of studies for each topic, and the total number of women. This information would help give an idea of whether there are likely to be sufficient data for the planned IPD analyses. The authors could also invite readers of the protocol to notify them of any missing studies.

There should be a section on inclusion and exclusion criteria. Whilst the detail could be cross referenced to their earlier work, it would help readers of this protocol if there was a summary of the criteria for each topic here. Also, what are the criteria for ‘inadequate data quality’ and ‘incomplete data’? One of the potential advantages of IPD is that exclusions due to missing data from aggregate data analysis can sometimes be reduced. Hence studies excluded from the earlier reviews for missing data should not be automatically excluded from this analysis.

Asking for the full dataset for each study sounds like the analysis will be a huge task. The usual strategy for IPD within trials (and if this is likely to be different from diagnostic tests it would be useful to have the rationale in the protocol) is to agree which variables will be used for the meta-analysis, and then ask for these only. This has the advantage that the collaborators identify the appropriate variables within their own datasets, and that data are anonymous. If the original hard copies are all that is available for some studies, will these be included?

It would clearly be useful to have the study protocols and CRFs, but in reality these are likely to be hard to come by. Hence ‘we will request’ would be more realistic than ‘we will obtain’.

For assessment of data quality it would be useful to have more detail. How will
completeness and quality be assessed and judged? For example, will there be range checks, checks for internal consistency, checks of computed variables against the original data? It is important the protocol is apriori clear about the criteria for exclusion based on quality. It might also be useful to plan a sensitivity analysis based on study quality.

P12 line 3 refers to merging studies ‘to form one extensive database’. This implies data analysis will not be within study – which is not the case for IPD. Although later sections refer to analyses within each study it should be clear throughout the protocol that data always remain within study, with pooling across studies. This paragraph also refers to ‘subgroups on all relevant issues concerning the clinical problems’. All such subgroups should be clearly specified within this protocol, as should the methods for testing interaction effects.

Much of the information on specific methods for the topics might be better presented in table format, or as lists. This would make it easier for readers to compare and contrast - patient characteristics and variables in the models, for example. For prediction of preterm birth, multiple pregnancy should be included in patient characteristics.

Prognostic studies are not covered in the protocol, and so should be dropped from the title and discussion. The title would be more accurate as ‘… prediction and diagnostic studies.’

Minor comments

The terms ‘preterm delivery’ and ‘preterm birth’ are used – it would be better to use just one of these, and preterm birth is preferable.

Page 8 in the section on preterm birth it is unclear whether the main outcome is preterm labour or preterm birth – the final sentence implies preterm labour. Although this is clarified later, it would be helpful to clarify here too.

Page 8, last para, first sentence starts by discussing couples with subfertility, and then relates this to the woman’s age. Is the man’s age not a factor too?!

Lelia Duley

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests