Reviewer's report

Title: Study protocol. A prospective cohort study of unselected primiparous women: the pregnancy outcome prediction study (POPS).

Version: 2 Date: 30 October 2008

Reviewer: John H Kalbfleisch

Reviewer's report:

Major Compulsory Revisions - None

Minor Essential Revisions - None

Discretionary Revisions and Comments for the Authors

(1)
The prospective cohort design is the fundamental design. After the database achieves a reasonable size, subsets of the database can be selected according to other study design guidelines. In these cases, it will be research conditions or constraints that dictate the study of a subset of the main database of information (the desire to control known bias-factors can be achieved with the case-cohort design and the high cost of predictor information can be accommodated with the case-control design using selected outcomes to define the cases). It would be helpful for some readers to have this clearly indicated in the analytic approach section. I would not want readers to interpret the overall project design as 3 parallel disjoint study designs.

(2)
Data analysis will search for different predictor measures (and combinations of them) that can be statistically related to adverse outcomes. The authors indicate the use of (high) sensitivity rates for the outcome (as if in a screening scenario). Specificity rates need to be concomitantly evaluated and reported (along with sensitivity). For selected outcomes, statistics of PPV and NPV can be then determined for different hypothetical test populations by allowing the outcome prevalence rate to vary over a range encountered in practice. This allows readers to anticipate false positive and false negative experience in patient populations where outcome prevalences will be low. In summary, specificity needs mention when sensitivity is presented.

(3)
Power calculations.

Statistical power is commonly used to justify or determine sample sizes for groups that are to be compared on some statistical measure (a mean level, group proportion, survival indicator, ..etc). The authors illustrate how sample size relates to assure statistical variability limits for diagnostic testing statistics.
(sensitivity). My suggestion would be to not use the concept of “power” in text since this is not the usual sample size – power dilemma that readers encounter. The authors might replace “Power calculations”, with “Recruitment Sample Size.”

(4)

Case-cohort studies of binary outcomes can use a survival analysis and logistic regression. The number of cases when time-to-event is censored will determine the appropriate analysis. Over-fitting of statistical models is likely to be unfamiliar to many readers (but the data analyst would be cognizant of the model fitting); the cross-validation discussion is appropriate and useful to readers. Predictability of screening is sometimes expressed in diagnostic testing terms of NPV, PPV (as indicted (2)above).

(5)

Serial data (within patients) will permit the study of trends and deviations from a usual pattern over the months up to delivery. These trends (not just values at one time point) can also serve as candidate predictors of adverse outcomes.

(6)

I wonder if it is possible to collect some data on cases electing not to participate in the study (certainly some demographic and baseline data exists). Would their adverse outcome experience be the same, better or worse? A similar comment can be made for women with parity higher than 1.

**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. John H. Kalbfleisch, Ph.D.