Reviewer's report

Title: Screening for inter-hospital differences in cesarean section rates in low-risk deliveries using administrative data. Can it contribute to quality of care initiatives?

Version: 2 Date: 26 April 2007

Reviewer: J C Glantz

Reviewers report:

Using national databases, the authors examine cesarean section rates at Belgian hospitals and correlate these with outcomes. Their goal is to determine whether cesarean rates outside (above or below) the normal range are associated with more or fewer adverse outcomes, and to use this information to report back to hospitals for quality assurance purposes. They did not specifically risk adjust in their analysis, but limited the study subjects to those they considered “low risk.” They conclude that cesarean section rates are rising, that there is significant interhospital variation in cesarean section rates, and that the information on the association of outcomes and cesarean section rates is inconsistent but may favor “average” rates.

Overall, this is a lengthy and dense article that is difficult to read, with a multitude of tables packed with numbers that are difficult to comprehend. I spent several hours reading and re-reading the manuscript, trying to understand the investigators’ results and their interpretation of these results. The introduction alone is 4 pages long, and the authors do not state their purpose until the final three sentences.

Major Compulsory Revisions

1. The first of my difficulties had to do with the lack of consistent definitions. The authors never actually define what they mean by elective cesarean (scheduled? patient choice? discretionary?) On page 2 of the introduction, they note that, “Gregory et al identified a concise set of clinical indications for elective (before labor) primary cesarean….,” but never say whether this is the definition used in their own study. Regarding their definition of “low risk,” it primarily appears to be vertex, term, singleton live fetuses, with weights between 2500 and 4500 grams. This is a rather general definition of low risk, and obviously fails to include some high risk patients such those with placenta previa, previous uterine surgery, and other such diagnoses that might mandate cesarean delivery. Because diagnoses such as abruption, preeclampsia, and diabetes also influence cesarean section rates, and because these diagnoses may be congregated in certain hospitals, it is unclear why the authors chose not to exclude some of these as well. The authors imply with some justification that there is some subjectivity in when a cesarean is indicted for these indications, but these diagnoses do provide a measure of patient risk status and have been used as such in other studies.

The authors never define what they mean by “semester:” From reading through the text and looking at the tables, I think they mean 6 month intervals, but they never state this specifically. Perhaps this is clear without definition in Belgium, but not necessarily in other countries.

I never really understood what their control and study populations were, because their comparisons primarily were amongst hospitals stratified by low-medium-high cesarean risk, not among subcategories of patients. It was not clear to me how they determined a significant departure in CSR as at least 35% above or 25% below the national period cesarean section rate (page 7) or the origin of their determination that a departure of 5% above or below the national trend was necessary in order to be classified as important. Without definitions such as these, the reader has an imperfect understanding from the start of what is being compared.

2. It is unclear why the authors chose to use one-minute Apgar scores as an outcome, considering that these are not thought to be clinically meaningful indicators of long term outcome.

3. I found the description of statistical methods difficult to follow. I understand why they used the fixed effects model, for example, but had trouble understanding their use of logistic regression, linear contrasts, rescaling techniques, and use of the generalized estimating equations. Perhaps someone with greater biostatistical expertise than myself should review the statistical methods.

4. In the first paragraph of the Results, what does “anamnesis of a previous cesarean delivery” mean, and why is this an important disagreement?

5. First paragraph, page 12, Results: when the authors refer to their study and control populations, are they
referring to low and high risk groups? Later in that paragraph they refer to outliers, but in the following sentence regarding trend, they state that no outlying hospitals were encountered. It is not clear what they mean by this. I have trouble understanding this entire paragraph, including the final several sentences referring to “hospitals departures, few cases left out, and departures of period and trend.”

6. One of the limitations of the study is the inability to adjust for socioeconomic indicators. With the multiple databases available to the authors, was there no information at all on such variables?

7. In the second paragraph of the Neonatal Outcomes portion, what do the authors mean when they state that “this finding suggests a lack of adequacy between selected mode of delivery and perinatal outcome?” In the third paragraph of this section, what do they mean when they say that, “the odds ratios…as well as an interaction term between CSR and type of delivery are mutually adjusted…”?

8. The Discussion is seven pages long and does not tie the paper together, nor does it leave the reader with a clear impression of the authors’ primary conclusions. It is quite long, and although it addresses some of the strengths and weaknesses of the study, contains considerable peripherally related information such as two pages about how such information might be of use to hospitals. Ultimately, reader will have problems understanding what the authors’ main points are. For example, a major issue in this study is whether a certain CSR produces optimal neonatal outcomes. Of the seven pages of discussion, only one paragraph (in the middle of the fifth page) gives any indication of the authors’ opinion, and only one sentence at that (“These considerations suggest that the ‘average CSR’ group might be the more adequate benchmark.”). The authors also imply in this paragraph that vaginal delivery is associated with better outcomes than delivery by cesarean section, but without further adjustment for medical and obstetrical diagnoses, such a statement is tentative at best.

9. There are many tables of information, most of which are very dense and hard to understand. Using table 2 as an example, I don’t recall the authors ever defining the difference between period and trend. I am not certain what they mean by the upper and lower bound of the Bonferroni confidence interval, although I assume it is an adjusted confidence interval that is greater than the 95%CI. The p values are carried out to 5 digits, which is far in excess of what they need to be. It is very difficult for the reader to understand what this table is trying to impart.

10. On table 3, are these odds ratios adjusted or relative to some reference within each group? What is the reference group? Tables 4 and 5 contain large amounts of data without comparison testing. Table 6 appears to be selective comparison testing based on the preceding two tables, but only for low one minute Apgar score. Table 7 contains adjusted odds ratios for a number of neonatal outcomes based on various delivery variables, but without reading every number, the reader has no sense of what is important in this table. Perhaps bolding or at least asterisks by significant odds ratios would help draw attention to significant values.

11. The duplex graphs in figures 1 and 2 appear virtually superimposable to the untrained eye. The lower portion of each graph (i.e., “congenital anomalies excluded”) could be deleted and mention made of this in the text.

By the end of this manuscript, I was left with the sense that there is some variation in cesarean section rates in “low-ish risk” women, but that it is difficult to generalize whether it is worse from an outcomes perspective to have a higher-than-expected rate or lower-than-expected rate. It is almost impossible for a study with limited neonatal follow-up data to answer that latter question, and much of what is coded is really surrogate information that may be associated with—but not equivalent to—adverse long-term outcomes (e.g. transient tachypnea does not predict chronic lung disease). The authors undertook a very complicated study to answer a multilayered series of questions, and clearly gave considerable thought to the study. They appropriately adjusted for some factors, used corrections for multiple comparisons, did not over-generalize their findings, and presented a large amount of data. They paid close attention to a myriad of details, but this same myriad of details risks losing the reader who may have difficulty seeing the forest for the trees. Perhaps this study would be easier to understand if it could be broken into smaller segments or written in such a way that emphasizes the authors’ main objectives and points in a clear manner. As it is currently written, however, it is dense and confusing, and the reader is left uncertain what to make of it.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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