Reviewer's report

Title: Geographic variation in cesarean delivery in the United States

Version: 2 Date: 8 April 2014

Reviewer: Jillian Patterson

Reviewer's report:

This is a well written paper adding to the literature on variation in cesarean section rates, considering geographic variation rather than variation at a hospital level. A large number of number of population and hospital market factors were considered in the analysis. The study aimed to identify potentially modifiable targets for intervention in order to reduce CS rates. The discussion however lacks 'next steps' in terms of how to apply this research to addressing rising CS rates.

1. Is the question posed by the authors well defined?

The authors address the issue of geographic variation in cesarean section rates, and explore whether drivers for cesarean section differ between privately insured and Medicaid patients, with the hope of identifying modifiable drivers of cesarean section.

2. Are the methods appropriate and well described?

The authors use a mix of logistic and hierarchical logistic regression models to assess the factors determining differences in CS rates between CBSA. The use of hierarchical regression is encouraging, as it accounts for correlations in the data, however the logistic regression models built ignore such correlations. More information on the particular hierarchical models used would be useful, particularly around random intercept vs random slope models, model fitting strategy and model assessment. The authors state that continuous variables were centered at 0, but what they mean by this is unclear- possibly that they centered variables at the population mean?

Given that other papers have found variation in CS rates by hospital (eg ref 10), why were hospitals not a level of the hierarchical model. Details on number of hospitals per CBSA and whether some hospital catchments spread across multiple CBSAs would be helpful for interpretation of results (particularly to a non-US audience).

Did the authors consider using a model which accounted for spatial correlation between the CBSA?

The results from the hierarchical model are difficult to interpret (large number of decimal points needed), particularly when presented as log odds of a 1% change. Perhaps the scale of the change reported could be changed (ie 5% or 10% change in population characteristic).

The reference to Figure 1 in the methods section should be moved to the results,
also please clarify whether this is based on the logistic regression or the hierarchical model. Is it possible to distinguish between CBSA/states which weren’t included in the analysis, and those which show no difference between Private and Medicaid rates?

3. Are the data sound?
Deliveries are identified by DRG, which may miss some deliveries where the initial hospital admission was for something other than delivery (Kuklina, Matern Child Health J, 2008). An assessment of the number of births missed, and births without mode of delivery information available would be useful. There is no indication of the validity of reporting of diagnoses or procedures used in defining the population or in adjustment. In the model, there is no adjustment made for parity or induction of labor, which could be expected to influence the cesarean section rate.

5. Are the discussion and conclusions well balanced and adequately supported by the data?
The authors report ‘notably’ more variation in CS rate for Medicaid compared with private insurance, and claim that these differences in rates can be partially explained by population and market characteristics, however there is little data given to back up this claim- including such information would be of value.

6. Are limitations of the work clearly stated?
The authors recognize the limitation of missing information on parity and other covariates, however do not discuss the potential impact of this on the results. Missing covariate information would be of concern if the distribution of parity and other covariates differed between Medicaid and private insurance, where it would appear as increased variation between CSBA.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

8. Do the title and abstract accurately convey what has been found?

9. Is the writing acceptable?
The writing, title and abstract are acceptable.

Major reviews
1. Please provide justification for the hierarchical models chosen, and the use of logistic regression to assess variation rather than the hierarchical model

Minor essential revisions
1. Please clarify the interpretation of log odds, and the relative importance of CSBA level factors compared with individual factors in contributing to observed variation

2. Move reference to Figure 1, and if possible remove CSBA not included in the
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