Reviewer's report

Title: A Retrospective cohort study of factors relating to the longitudinal change in birth weight

Version: 1  Date: 12 August 2015

Reviewer: Russell Kirby

Reviewer's report:

In this manuscript, the authors use a single center perinatal database to examine trends in birth weight among term singleton deliveries from 1995-2009. The study also examines trends in birth length and ponderal index. While the latter adds new information, most of the covariates for the analysis of birth weight could be obtained from population-based sources.

Major compulsory revisions

1) The authors do not explain why the study ends in 2009, given that the manuscript was submitted in 2015. There could be arguments supporting the choice to end in 2009, including perhaps that the hospital implemented a no unindicated intervention policy, but the authors need to substantiate this if it is true.

2) There is no attempt to explain how generalizable these data may be. How has the case-mix of the hospital obstetrical cases changed over time? How representative of the Cleveland metropolitan area are these data, and how if at all has this changed over the study period?

3) Where are the results of the separate regressions by race/ethnicity (p 5 line 105-6)?

4) In table 1, what statistical test was used to determine significance as shown in the last column? Did the authors conduct a test for trend, either by 3-year interval or single year? There are sufficient data to conduct JoinPoint regression for single year - this test determines if there is a significant trend, and also if there are one or more inflection points to the trend. There are other potential tests as well.

5) It is hard to make sense of the regression results provided in Tables 2-4. What data are provided in the first column - headed 'standard estimate (correlation coefficient)'? Is this in fact 'r'? If so, it would be better to label the column thusly. It would also be helpful to show the t-statistic for each variable, as this gives the reader a sense of the relative contribution of each variable. It would seem, and unsurprisingly, that gestational age is the biggest contributor in each model. Also, give a measure of goodness of fit for each model as a footnote below.

6) In Figure 1, be consistent in terms of capitalization - shouldn't 'time' be
capitalized in the title for each panel? And what is 'mean' - its really not that, it is the overall without respect to the subgroups shown in each panel. What is 'wga'? This is an unfortunate abbreviation - better to say 37 weeks, 38 weeks, etc, and indicate its gestational age in the panel title.

Continuing, how were the trend lines fitted in each panel of Figure 1? This does not appear to have been mentioned in the methods.

7) in discussing limitations, no mention is made of the fact that this is a single-center study and its rather difficult to assess its broader generalizability.

8) Given the comments about Zhang's work, why did the authors not examine also the patterns among spontaneously induced deliveries? Surely these can be identified within the database?

9) Given the almost non-findings regarding the Ponderal Index (over R-sq of 0.05, corr coeff as indicated in Table 4 of 0.11), the statements about its utility (p 9, 203-210 as well as comments in conclusion) seem unwarranted.

10) There is a much broader literature on this subject, including now studies examining maternal weight growth velocity during pregnancy, and also work examining effects of weight retention between pregnancies on perinatal outcome in subsequent pregnancy. It would be interesting for the authors to comment on this in terms of implications for future work, and also to see if the there are sufficient subjects in the sample with two consecutive pregnancies to conduct a substudy on this topic.

Minor essential revisions

1) p 5, line 99 and elsewhere: this would read better as 'as well as in three-year intervals' - epoch is usually reserved for much larger periods of time

2) place each table on a separate page. Add information about what statistical tests were performed.

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.