Author’s response to reviews

Title: Derivation and validation of a model predicting the likelihood of vaginal birth with labor induction

Authors:

Sepand Alavifard (s.alavifard@mail.utoronto.ca)
Kennedy Meier (kennedy.meier@mail.utoronto.ca)
Yonatan Shulman (shulmanyonatan@gmail.com)
George Tomlinson (george.tomlinson@thebru.ca)
Rohan D’Souza (Rohan.DSouza@sinahealthsystem.ca)

Version: 1 Date: 12 Jan 2019

Author’s response to reviews:

José Guida, MsC, M.D. (Reviewer 1)

Dear authors, thanks for the opportunity of reviewing this interesting manuscript. Performing IOL is becoming more and more common on daily practice and having a tool to predict success of vaginal birth is a key tool to convince women in their decision. In countries where c-section rates are raising year-by-year, a tool like this is essential.

(1) However, after reading your manuscript, my feeling is that I could not understand clearly how each of the variables impact on the likelihood of vaginal birth. Of course after reading table 2 it became clear, however to someone not used with statistical it may be hard to understand. I suggest that you include a table or even a figure to make easier to understand the impact of each variable on your model.

Authors’ response: We thank Dr. Guida for pointing out that the impact of each of variable on the likelihood of vaginal delivery was not well presented. We do agree that using figures would help present the impact of variables on the outcome. In order to do this for a multivariable regression model, all variables other than the one being presented have to be static. Therefore, we have used the values from Table 1 (average values in the study population) to develop Figure 2, below. Specifically, it is possible for cervical dilation, maternal age and gestational age to take on most combinations of biologically plausible values, so it is reasonable to interpret the graphs in Figure 2 as the effects of each variable with the others (and others not showing in these figures) held constant at their averages. However, the three weight-related variables (pre-pregnancy weight and BMI and delivery weight) are highly collinear. This does not pose a problem for the prediction model, as predictions are made at these correlated values of weight and BMI. We would be reluctant, for example, to use our prediction model on a very heavy...
woman with a very low BMI as this observation would lie outside the range of data on which the model was fitted (and would therefore be an extrapolation to a case where we cannot be confident in our model). However, collinearity of the predictors does hamper the interpretation of the coefficients (or the related polynomial curves), as, for example, we would have to think about what it would mean to examine the effect of BMI changing from low to high values at a fixed value of pre and delivery weight. For any fixed weights, BMI in the sample takes on only a restricted set of values. Similarly, at the mean pre-pregnancy weights, there is a relatively narrow range of possible delivery weights. It is the trio of values (weight-pre, BMI-pre and weight at delivery) that move together that need to be considered in obtaining predictions.

The following changes have been made to the manuscript

Results – Model Development: The relationship between the likelihood of vaginal delivery with each of maternal age, gestational age, cervical dilatation, and parity are illustrated in Figure 2; for each figure, the value for all other variables are held static as per the study population’s average values (Table 1). The relationship between the outcome and pre-pregnancy weight, pre-pregnancy BMI, and weight at delivery was not illustrated due to collinearity. For any fixed weights, BMI in the sample takes on only a restricted set of values. Similarly, at the mean pre-pregnancy weights, there is a relatively narrow range of possible delivery weights. It is the trio of values (weight-pre, BMI-pre and weight at delivery), which move together that need to be considered in obtaining predictions.

A new figure has been added: Figure 2 – The relationship between maternal age (A), gestational age (B), cervical dilatation (B), and likelihood of vaginal delivery stratified by parity, nulliparous (black-line) vs. multiparous (grey-line).

It is important to note that we considered providing the OR associated with each variable; however, given that our final model has polynomial variables, the odds ratio (OR) for each variable does not have a clinical relevance and nor can a reader look at the combination of ORs for a polynomial term (e.g. weight, weight2, and weight3) and grasp a variable’s overall impact on the outcome. The below table is able to illustrate our point. For this reason, we had opted not to include ORs in Table 2. If the editors feel strongly about including the OR (95% Confidence Intervals) for each term, we could modify Table 2 to include this information.

<table>
<thead>
<tr>
<th>Term</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity (Multiparous)</td>
<td>6.02</td>
</tr>
<tr>
<td>Weight, pre-pregnancy (kg)</td>
<td>3.85E+26</td>
</tr>
<tr>
<td>(Weight, pre-pregnancy)2</td>
<td>3.94E-06</td>
</tr>
</tbody>
</table>
(Weight, pre-pregnancy)$^3$ 0.0015

BMI, pre-pregnancy (kg/m²) 6.63E-26

(BMI, pre-pregnancy)$^2$ 12354.11

(BMI, pre-pregnancy)$^3$ 322591.68

Gestational Age (weeks) 0.0024

(Gestational Age)$^2$ 0.034

(Gestational Age)$^3$ 92.22

Weight, at-delivery (kg) 8.54E-07

(Weight, at-delivery)$^2$ 2412.92

Dilation (cm) 1.72E+10

(Dilation)$^2$ 143.60

Maternal Age (years) 0.00022

(Maternal Age)$^2$ 0.013

(2) I also think that the final equation of your model should be published in this manuscript (of course I understand you will use it in another article, probably performing external validity).

Authors’ response: We thank Dr. Guida for commenting on including the final equation of the model in the manuscript. Although Table 2 presents the final equation as components, we agree that this can be unclear for readers. Therefore, we have added a row to Table 2 explicitly presenting the equation.

(3) Other suggestion is to add some more information about the Akaike Information Criterion on your methods section.

Authors’ response: We thank Dr. Guida for suggesting additional information to be added about the AIC. We have modified the sentence to say “Predictors were then excluded by backward-elimination based on the Akaike Information Criterion (AIC), a criterion approach based on the residual sum of squares, used to compare non-nested models to identify an optimal subset of predictors.”
(4) Also, the first phrase of the 1st paragraph of your discussion is out of context.

Authors’ response: We agree with Dr. Guida’s comment on the first phrase of the manuscript’s discussion, and have removed it.

(5) Finally, figure 3 has to be improved because its resolution is poor in current format and also I would add Bishop data on patient 1 and 3 (I know they are the same, however it would be more informative on your figure to have all data).

Authors’ response: We thank Dr. Guida the comments on Figure 3 (now Figure 4). As recommended, we have attached added data on the Bishop score for patient 1 and 3 and attached a higher resolution version.

Thanks again for this opportunity and I really hope this model be very successful helping clinicians worldwide.

Olimpio Moraes (Reviewer 2)

I congratulate the authors for the paper. This instrument with its external validation will be of great importance for obstetrics and the valorization of bioethical principles in the decision making of the induction of labor or cesarean section.

Leila Katz (Reviewer 3)

Regarding the Manuscript PRCH-D-18-01262, first of all, I would like to thank you for the opportunity to review the manuscript. It is very well written. It is very relevant, includes a very interesting population and methods are well carried out. However there some points that should come to attention:

(6) In the abstract, in line 35, you state that a proportion of inductions end in emergency cesarean deliveries. This is true, but some of them also end in c-sections because of failed induction, or because long inductions may lead doctors and patients to decide for ending inductions and opting for cesarean. So the sentence would be more exact removing the Word "emergency".

Authors’ response: We agree with Dr. Katz’s recommendation and have removed the word “emergency”.

(7) There is a limitation and that should be more detailed in the discussion session. This is the fact that the only variable of the Bishop Score used in the model was cervical dilation.

Authors’ response: We agree with Dr. Katz’s recommendation and acknowledge that this was not adequately addressed in the discussion. We have added both suggested references and the following sentences to the discussion to better discuss this limitation: “It is possible that although variables such as cervical consistency and position may be less likely to influence the prediction of the success of vaginal birth, cervical effacement and fetal station might still be important predictors.28-32 We acknowledge that this is a major limitation that could compromise the model, and will consider these variables during the prospective global validation of this model.”

(8) In the discussion section, in lines 244-246, when you described the inclusion only of the cervical dilation in the model, it seems this was an option and not something that occurred because these characteristics were not available.

Authors’ response: We agree with Dr. Katz’s comment on how these lines can be misinterpreted. We have removed it from here, as a discussion on the inclusion of cervical dilatation alone in our model, fits better in the limitations section.

(9) In lines 266-268 this limitation is pointed out, but very superficially. Since this is something that may compromise the model, we feel that should be more detailed.

Authors’ response: Once again, we agree with Dr. Katz’s comment that this is a major limitation and could compromise the model. Unfortunately since we used retrospective data, we would not be able to reduce the missingness unless we made assumptions or used imputation methods, both of which could also compromise the model. The current model, despite including only one cervical variable still outperformed all other published models, and we hope that with the inclusion of other cervical variables on prospective validation, the predictive ability of the model will increase even further. We have clearly acknowledged this as a limitation and stated that we will include these variables during prospective validation of the model.

(10) Were all the cesareans, emergency cesareans? Were there no failed inductions? All women achieved active labor? Were there no other indications for c-sections other from emergency situations? This could be detailed in the result section and the flowchart.

Authors’ response: We thank Dr. Katz for this comment. As suggested, these details have been presented in the flowchart and in the results section as follows: Most caesarean deliveries were performed for concerns with fetal heart rate patterns in labour (136/290) and arrest of dilatation.
in the active phase of the first stage of labour (102/290), and only a small minority (16/290) for failure to achieve active labour (Figure 1).