Received: 13 May 2020 | Revised: 11 August 2020 | Accepted: 29 September 2020

DOI: 10.1111/aogs.14020

RESEARCH ARTICLE

Predicting vaginal birth after previous cesarean: Using machine-learning models and a population-based cohort in Sweden

Charlotte Lindblad Wollmann[1,2] | Kyle D. Hart[3] | Can Liu[1] | Aaron B. Caughey[3] | Olof Stephansson[1,2] | Jonathan M. Snowden[3,4]

[1]Clinical Epidemiology Division, Department of Medicine, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden

[2]Division of Obstetrics and Gynecology, Department of Women's and Children´s Health, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden

[3]Department of Obstetrics and Gynecology, Oregon Health & Science University, Portland, Oregon, USA

[4]School of Public Health, Oregon Health & Science University and Portland State University, Portland, Oregon, USA

Correspondence
Charlotte Lindblad Wollmann, Clinical Epidemiology Unit, T2, Karolinska University Hospital, Solna, 17176 Stockholm, Sweden.
Email: charlotte.lindblad-wollmann@sll.se

Funding information This study was supported by grants from the Swedish Research Council (2013-2429, OS) and by grants provided by the Stockholm County Council (ALF project 2013-0156, OS).

Abstract

Introduction: Predicting a woman's probability of vaginal birth after cesarean could facilitate the antenatal decision-making process. Having a previous vaginal birth strongly predicts vaginal birth after cesarean. Delivery outcome in women with only a cesarean delivery is more unpredictable. Therefore, to better predict vaginal birth in women with only one prior cesarean delivery and no vaginal deliveries would greatly benefit clinical practice and fill a key evidence gap in research. Our aim was to predict vaginal birth in women with one prior cesarean and no vaginal deliveries using machine-learning methods, and compare with a US prediction model and its further developed model for a Swedish setting.

Material and methods: A population-based cohort study with a cohort of 3116 women with only one prior birth, a cesarean, and a subsequent trial of labor during 2008-2014 in the Stockholm-Gotland region, Sweden. Three machine-learning methods (conditional inference tree, conditional random forest and lasso binary regression) were used to predict vaginal birth after cesarean among women with one previous birth. Performance of the new models was compared with two existing models developed by Grobman et al (USA) and Fagerberg et al (Sweden). Our main outcome measures were area under the receiver-operating curve (AUROC), overall accuracy, sensitivity and specificity of prediction of vaginal birth after previous cesarean delivery.

Results: The AUROC ranged from 0.61 to 0.69 for all models, sensitivity was above 91% and specificity below 22%. The majority of women with an unplanned repeat cesarean had a predicted probability of vaginal birth after cesarean >60%.

Conclusions: Both classical regression models and machine-learning models had a high sensitivity in predicting vaginal birth after cesarean in women without a previous vaginal delivery. The majority of women with an unplanned repeat cesarean delivery were predicted to succeed with a vaginal birth (ie specificity was low). Additional

Abbreviations: AUROC, area under the receiver-operating characteristics curve; CI, , confidence interval; CD, cesarean delivery; TOLAC, trial of labor after cesarean; VBAC, vaginal birth after cesarean.

covariates combined with machine-learning techniques did not outperform classical regression models in this study.

KEYWORDS

Cesarean delivery, machine-learning, prediction, random forest, trial of labor, vaginal birth after cesarean

1 | INTRODUCTION

The choice between a trial of labor after cesarean (TOLAC) and an elective repeat cesarean delivery (CD) may be challenging. A successful TOLAC, a vaginal birth after cesarean (VBAC), decreases epidemic CD rates and maternal morbidity associated with multiple CDs.[1,2] Yet, TOLAC bear the risk of uterine rupture or an unplanned repeat CD and increased the risk of adverse outcomes,[2,3] although generally being considered safe and encouraged in many countries.[4]

Success rates of TOLAC vary between 40% and 80% internationally.[3,5-7] Analyses using decision models concludes, based on risks associated with VBAC, that when the chance of VBAC is greater than 50%-70%, TOLAC should be offered.[6,8-10] However, TOLAC rates vary depending on individual women's preferences.[11] Therefore, predicting individual probability of VBAC could facilitate the decision-making.

Grobman et al developed a model for predicting VBAC based on multivariable logistic regression,[12,13] further modified and evaluated in the Swedish setting by Fagerberg et al.[14] Both Grobman and Fagerberg included women with previous vaginal delivery, one of the strongest predictors for VBAC.[6,15] However, no model has previously been developed for women without previous vaginal delivery, whose outcomes are more unpredictable for clinicians.

With the growing availability of data, machine-learning methods might have an advantage as prediction tools in healthcare,[16-18] with the ability to consider many candidate predictors, taking into account complex relations (eg, complex interactions, non-linearity).[16,19,20] These algorithms sometimes include surprising predictors that human investigators might not otherwise have considered.[21,22] The results may improve clinical counseling, if accuracy is high.[23,24]

Improving quality of care and counseling and better predicting vaginal birth in women with only one prior cesarean and no vaginal deliveries would greatly benefit clinical practice and fill a gap in research. Our primary aim was to develop individualized pre-delivery prediction models for VBAC using conditional inference tree, conditional random forest and lasso binary regression. We built on a prior study where women with a first unplanned cesarean were associated with a higher risk of repeat CD compared with women with elective first CD. However, almost 70% of all women eligible for TOLAC had a vaginal birth.[5] Recognizing that prior vaginal birth strongly predicts VBAC, we focused on predicting VBAC among women with only one prior birth, a cesarean, since prediction in these women is a great challenge in the clinics. Our second aim was to compare our models with previous classical regression models.[12-14]

Key message

The majority of women with an unplanned repeat cesarean delivery were predicted to succeed with a vaginal birth (ie specificity was low). Additional covariates combined with machine-learning techniques did not outperform classical regression models in predicting vaginal birth after cesarean.

2 | MATERIAL AND METHODS

2.1 | Source of data

Prospectively collected data on maternal, delivery and infant characteristics were obtained from the population-based regional Stockholm-Gotland Obstetric Cohort.[25] The cohort includes all singleton births (n = 175 522) between January 2008 and October 2014 at seven hospitals in the region. Approximately 25% of all annual births in Sweden occur in this region. Almost all pregnant women in Sweden utilize standardized antenatal care, offered free of charge. The cohort is based on daily, automatically forwarded data from the electronic medical record system (Obstetrix, Cerner Inc.) used at all antenatal, ultrasound, delivery and postnatal care units in the region. Maternal and infant information from prenatal care, delivery and the postpartum period are prospectively entered into the medical records by midwives and physicians in a standardized way.[25]

2.2 | Participants

We extracted information on women with a first CD and a second singleton delivery during the study period 2008-2014. We restricted the second delivery to liveborn infants in cephalic presentation at 37 gestational weeks. Of the 5302 women with a first CD and a subsequent delivery in the Stockholm-Gotland Obstetric Cohort, 41% had an elective repeat CD and were excluded from our study, leaving 3116 women performing a TOLAC (Figure S1). Further details on data collection methods and features of this study population of 3116 women with a TOLAC are available elsewhere.[5]

2.3 | Outcome

Our primary outcome was to study the performance of three machine-learning methods regarding the ability to predict probability of VBAC (area under the receiver-operating characteristics curve (AUROC), accuracy, sensitivity and specificity). The secondary outcome was to compare the predicting performances with a well-used prediction model from the USA (Grobman et al)[12] and a Swedish version of the Grobman model (Fagerberg et al).[14]

2.4 | Predictors

Our intention was to inform clinical counseling before labor onset, so we set the temporal point of prediction before 37 gestational weeks, prior to term labor onset. We considered data from the first antenatal visit and all subsequent visits, before data about the second delivery were known. We included maternal characteristics from both the first and the second pregnancy, variables related to the first pregnancy and CD, and information about first infant, pre-gestational health conditions, conditions that developed during either pregnancy, and information on each maternity hospital (ie all factors presented in Table 1). We also included sex of the second infant. Intended onset of second labor was included, since this is important for a successful TOLAC (Table 1).[5,6]

2.5 | Overall statistical approach

We divided the study population into a training (n = 1558 women) and a validation (n = 1558 women) set, using a 1:1 split by random sampling, and predicted VBAC in the validation dataset using the estimates reported by Grobman and Fagerberg. We then fit new logistic regression models using the same specification as the Grobman and Fagerberg models using the training dataset, and summarized their performance in the validation dataset. Finally, we fit a conditional inference tree,[26] a conditional random forest[27] and a lasso binary regression model using the training dataset and summarize their performance in the validation dataset, and[28] compared the predictive performance of each.

As a sensitivity analysis, we fit the new models on the complete (training+validation) dataset and estimated classification error using fivefold cross-validation.[20]

2.6 | Data management

A detailed overview of our data management and missing data approach is provided in Appendix S1 and Table S1.

2.7 | Statistical analyses

We divided the study population (n = 3116 observations, with all of the applicable variables used in the Grobman[12] and Fagerberg[14] models, in addition to the predictors described above) into a training (n = 1558 women) and a validation (n = 1558 women) set, using a 1:1 split by random sampling. Missing data were replaced using

TABLE 1 List of candidate predictors, Study population from the Stockholm-Gotland Obstetric Cohort, 2008-2014

Variables related to pregnancy and infant #1	Variables related to pregnancy and infant #2
Maternal	Maternal
Mother's height	Mother's height
Family situation	Mother's age
Pregnancy	Mother's BMI
In vitro fertilization	Change in BMI (between first antenatal visit in pregnancy with Infant 1 and Infant 2)
Successful external cephalic version	Family situation
Any hypertensive disorder	Tobacco use (in either pregnancy)
Delivery	Pregnancy
Onset of labor[a]	Pregnancies between infants (including second infant)
Medical induction	Inter-pregnancy interval (years)
Mechanical induction	In vitro fertilization
Cervical dilation before CD	Any hypertensive disorder
Fully dilated cervix before CD	Delivery
Recurrent CD indication	Hospital rate of elective repeat CDs
CD indication[b]	Hospital rate of unplanned CDs
Hierarchical indication for 1st CD[c]	Onset of labor (induction, spontaneous)
Blood loss volume	Characteristics of infant
Puerperal or postpartum infection	Neonate sex
Maternal length of stay in hospital	Gestational age
Characteristics of infant	
Neonate sex	Variables related to either pregnancy, maternal disease
Gestational age (GA)	Lung disease
GA-standardized birthweight	Psychiatric or psychological disorder
Head circumference (cm)	Endocrine disease
APGAR 1 min	Recurrent urinary tract infections
APGAR 5 min	Gynecological disease
APGAR 10 min	

[a]For infant #1: Planned CD, induction, or spontaneous For infant #2: induction or spontaneous.

[b]Dystocia, non-reassuring status, elective, other.

[c]As defined by Carlsson Wallin et al (30).

single imputation (Table S2). To test the original Grobman and Fagerberg models in our dataset, we used the originally reported log odds as offsets in a logistic regression model to predict VBAC in the validation dataset. We then refit both models in the training dataset and repeated prediction in the validation dataset. We omitted race and ethnicity variables from our implementations of the Grobman and Fagerberg models for several reasons: Race and ethnicity data were unavailable in our dataset; these variables have a different implication in a Swedish population than in the American population where the Grobman model was developed;[29] and there is increasing awareness that inclusion of race in prediction models is often unwarranted on theoretical grounds.[30] Because our population comprised only women with a first CD and a second TOLAC, we also excluded variables for prior vaginal delivery and prior VBAC.

We trained a conditional inference tree,[26] a conditional random forest[27,31] and a lasso logistic regression model[20,32] in the training dataset and then used these models to predict VBAC in the validation dataset. For all models, we calculated AUROC, accuracy, sensitivity and specificity in the validation dataset, based on a 50% decision cut-off for predicted probability. We constructed calibration curves from the validation dataset for each model by coarsening predicted probabilities into bins of 0.05 width and calculating the proportion of observed VBACs within each bin. The calibration curves compares predicted to observed probability of VBAC and provides a view of model performance across the range of predicted probability. For all new models, we also estimated classification error using fivefold cross-validation in the entire (n = 3116) dataset.

The conditional random forest was grown to 200 trees, based on examination of out-of-bag error, with the *m* parameter set to 7 (of 42 candidate predictors). We tuned the lambda parameter for the lasso model by selecting the value of lambda associated with the smallest error in the fivefold cross-validation.

All data management and statistical analyses were performed using R version 3.5.1. The conditional inference tree and conditional random forest were grown using the party package version 1.3-1. The lasso model was fitted using the glmnet package version 2.0-16.

2.8 | Ethical approval

The regional ethical committee at Karolinska Institutet, Stockholm, Sweden, approved the study protocol (No 2009/275-31, approved 2 April 2009).

3 | RESULTS

Of all participating 3116 women performing a TOLAC, 69% (n = 2146) had a vaginal birth and 31% (n = 970) a repeat CD (Figure S1).

Table S3 describes the characteristics of the participants by delivery mode in second delivery. Compared with women with a repeat CD, women who had VBAC were more likely to be younger, taller, have a lower body mass index (BMI) and a lower change in BMI from first to second pregnancy. They were more likely to have spontaneous labor onset and deliver in a hospital with lower rate of unplanned CDs in the second delivery. They were less likely to have labor dystocia as the indication of the first CD and to have been induced in the first delivery, and more likely to have reached second stage of labor before the CD or have an elective indication for the first CD.

Distributions of outcome and prediction variables in the training and validation datasets (Table S4) were very similar. Although gestational age in the second infant was statistically significant, the difference is not clinically meaningful.

Estimates for variables refitted in our data were similar to the estimates reported by Grobman and Fagerberg, with the exception of the hierarchical CD indication of preterm birth (defined as birth before 37^{+0} gestational weeks), which flipped direction from the model that Fagerberg reported (Table S5).

TABLE 2 Predictive performance of existing and new predictive models (95% CI)

Model	AUROC	Accuracy	Sensitivity	Specificity	Fivefold CV accuracy
Grobman (original estimates)	0.64 (0.61-0.67)	69.9% (67.6%-72.2%)	97.6% (96.7%-98.5%)	7.1% (4.8%-9.4%)	NA
Grobman (refit model)	0.64 (0.61-0.67)	69.9% (67.6%-72.2%)	96.5% (95.4%-97.6%)	9.6% (7.0%-12.3%)	69.0% (67.4%-70.7%)
Fagerberg (original estimates)	0.63 (0.60-0.66)	70.1% (67.8%-72.4%)	91.6% (89.9%-93.2%)	21.4% (17.7%-25.1%)	NA
Fagerberg (refit model)	0.66 (0.63-0.69)	70.7% (68.5%-73.0%)	93.2% (91.8%-94.7%)	19.7% (16.1%-23.3%)	70.1% (68.5%-71.7%)
Conditional inference tree	0.61 (0.58-0.63)	69.4% (67.1%-71.7%)	100.0% (100.0%-100.0%)	0.0% (0.0%-0.0%)	68.4% (66.8%-70.0%)
Random forest	0.69 (0.66-0.72)	70.0% (67.8%-72.3%)	97.9% (97.0%-98.7%)	6.9% (4.6%-9.2%)	69.9% (68.3%-71.5%)
Lasso	0.67 (0.64-0.70)	70.4% (68.1%-72.7%)	93.4% (92.0%-94.9%)	18.2% (14.8%-21.7%)	70.4% (68.8%-72.0%)

Abbreviations: AUROC, area under the receiver-operating characteristics curve; CV, cross-validation.

AUROC ranged from 0.61 to 0.69, with sensitivity (probability of correctly identifying a VBAC for second delivery) above 91% and specificity (probability of correctly identifying a repeat CD for second delivery) below 22% for all models (Table 2). The conditional inference tree assigned >50% probability of VBAC to every individual in the validation sample, giving a 100% sensitivity and 0% specificity. Specificity was poor in all models but was highest in the Fagerberg model (19.7%), while still maintaining sensitivity above 90% (Table 2). Accuracy (correctly classified delivery modes) ranged from 68.4% to 70.4%, and fivefold cross-validation accuracy was similar.

Accuracy, sensitivity and specificity were calculated by assigning a predicted outcome based on a probability cut-off of 50%. An alternative way to look at predictive accuracy is to compare the distribution of observed VBACs over the range of predicted probabilities; to that end, we have presented calibration plots for each model (Figure 1). In these calibration plots, all models except the random forest deviated from observed CD rates in the lower range of predicted probability (<50%) and all models had wide confidence bands in this lower range. In the Grobman and Fagerberg models, 53% and 73% of individuals with an unplanned repeat CD had predicted probability of VBAC above 60%; in the conditional inference tree, random forest, and lasso models, 97%, 61% and 60% of unplanned repeat CDs had predicted probabilities of VBAC above 60% (Figure 2).

The conditional inference tree selected splits at the indication for the first CD and the presence of any hypertensive disorder during the second pregnancy (Figure S2).

Variables with the highest conditional importance in the random forest included indication for the first CD, onset of labor for the first infant and maternal characteristics (Figure S3).

The lasso model also selected indication for the first CD, although the strongest predictor of VBAC in this model was being a single mother (vs cohabiting) (Table S6).

4 | DISCUSSION

On a population-based sample of women without previous vaginal delivery performing a TOLAC, we compared two existing prediction models with three new machine-learning models. AUROC was <0.70 for all models, sensitivity was >91%, and specificity was <22%. The majority of women with an unplanned repeat CD had predicted probability of VBAC >60%.

All AUROCs were slightly lower than in the original studies by Grobman (0.75)[12] and Fagerberg (0.74),[14] although those models included women with previous vaginal birth, a strong predictor of VBAC.[6,12,15] We excluded women with previous vaginal delivery, likely making prediction more difficult.

The indication for the first CD was identified as an important variable in all of the machine-learning approaches. This consistency lends credibility to the variable's use as a decision-making metric and further strengthens the notion that healthcare providers should emphasize improved birth outcomes in first-time mothers.

Our study is mainly restricted by the relatively short study period (2008-2014) and the limited hospitals included, which also constrained the sample size and the inter-pregnancy interval of women included. The limited sample size reduced the fidelity of the hierarchical classification of indication of CD in first delivery as used in the Fagerberg model. Our decision to use single, rather than

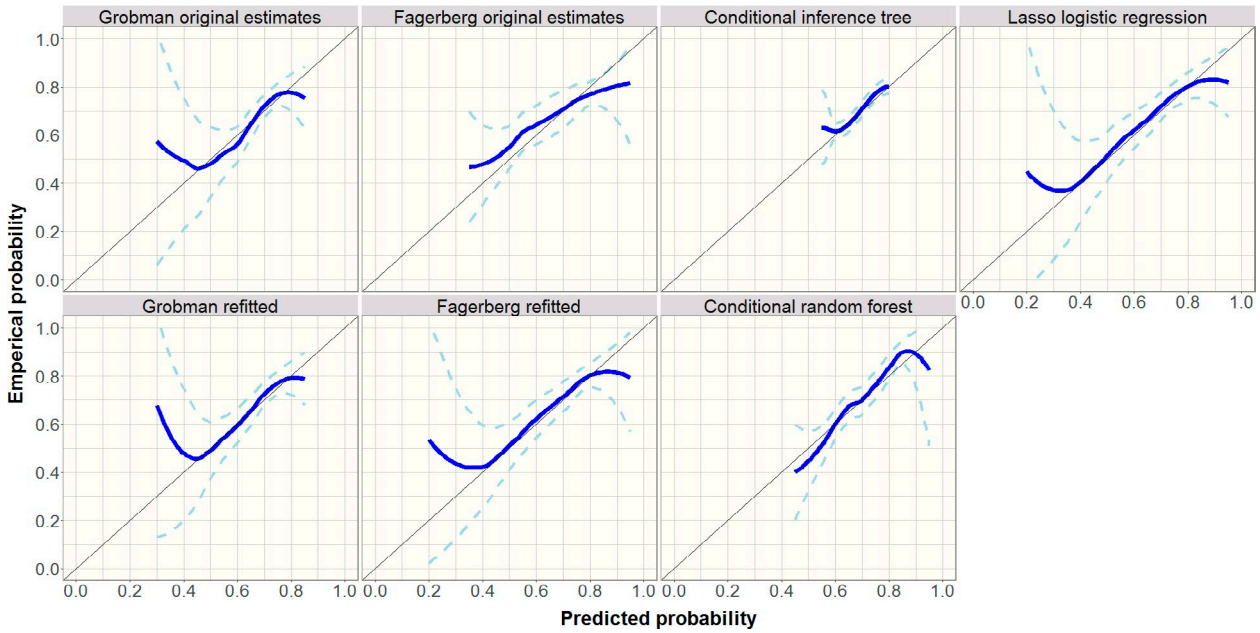

FIGURE 1 Calibration plots of the different prediction models. The solid blue line represents the actual performance with dotted 95% confidence bands. Solid gray line is the ideal performance [Color figure can be viewed at wileyonlinelibrary.com]

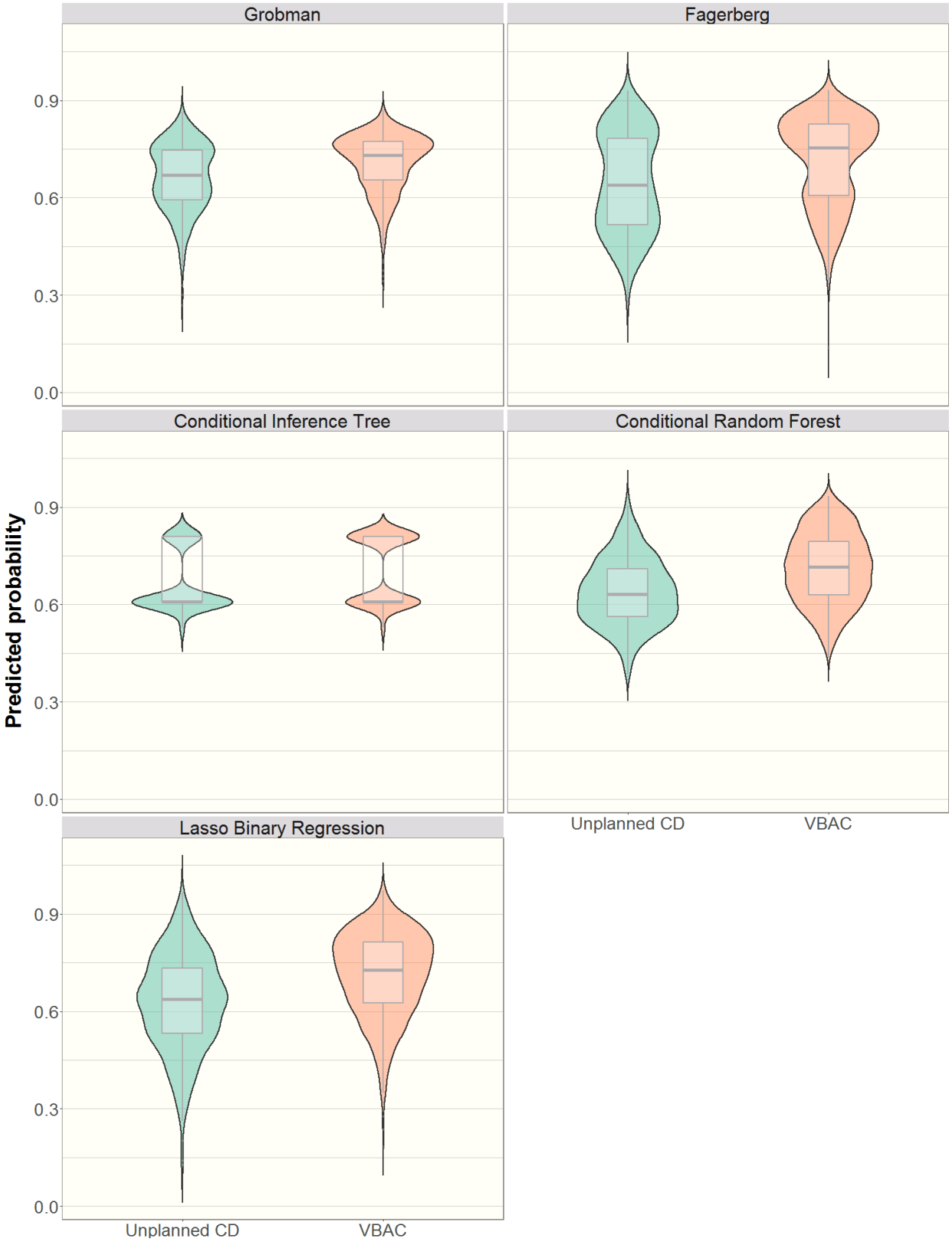

FIGURE 2 Distributions of predicted probability by observed VBAC status for existing and new models [Color figure can be viewed at wileyonlinelibrary.com]

multiple imputation likely resulted in higher variance in the imputed values, but a sensitivity analysis indicated that imputation did not substantially alter our results. Considering the number of predictors available for training, the conditional inference tree seems relatively short, and it produced a short range of predicted probabilities. The relatively poor performance of the tree may result from the high

variance associated with the method relative to it ensemble counterparts, such as random forests.[20]

Our population-based cohort with granular details based on the prospectively collected electronic medical records provides an array of clinically relevant maternal characteristics that were not used in previous models. With the growing use of the electronic medical records in many other countries, we believe our approach is transferable to other contexts where medical records are digitalized. This study is conducted in a context with universal maternity care, small variation of quality of care between hospitals and a high rate of TOLAC. This relatively equal opportunity for women to have a TOLAC makes the prediction model more representative of the chance of achieving VBAC among women with one previous CD, not affected by the substantial selection that occurs in settings with much lower TOLAC rates.

Despite the fine-grained dataset containing prospectively collected data, our models did not perform appreciably better than previous classical models, indicating that there may also be higher-level factors affecting TOLAC success on a patient, health provider, hospital and country level, as previous literature has suggested.[33-35] Future research should explore these factors (including obstetrician traits, maternal preferences, maternity unit staffing and workload during delivery), which might improve prediction of VBAC success.

In their current stage, none of the prediction models is very useful for women without previous vaginal delivery. Most women were predicted to have a vaginal birth. However, in all models, the majority of individuals with an unplanned repeat CD had predicted probability of VBAC >60%, which undermines the utility of even the better-performing top half of the models. Although increasing prediction accuracy by small amounts may have a limited effect on an individual patient's decision-making, it should be noted that even small increases in accuracy and increased VBAC success could have large effects at the population level, given the current drive to reduce cesarean use.

Sweden has a generous policy for TOLAC, encouraging women with one previous CD to have a TOLAC unless there is a medical contraindication for undergoing vaginal delivery (eg placenta previa). In this context of broad TOLAC access, improving prediction of repeat CD (specificity) should be the aim of future models. Considering that available maternity care services vary significantly between population groups, hospitals and countries, more restrictive policies may be common. In contexts where selected women are given the opportunity to try TOLAC, models with a good prediction of VBAC (sensitivity) may better fit the clinical purpose, encouraging women and health providers to provide a TOLAC more generally.

5 | CONCLUSION

It remains difficult to predict vaginal birth in women with only a prior cesarean. Both previous known models based on classical regression and new machine-learning models had a high sensitivity in predicting vaginal birth, with most women predicted to have a vaginal birth. However, the majority of women with an unplanned repeat CD were also predicted to succeed with a vaginal birth. Additional covariates combined with machine-learning techniques did not increase the prediction performance. There are most likely other factors affecting TOLAC success in a patient, eg factors on the hospital level, which may be subject to further research.

CONFLICT OF INTEREST

The authors have stated explicitly that there are no conflicts of interest in connection with this article.

ORCID

Charlotte Lindblad Wollmann https://orcid.org/0000-0003-0650-312X
Olof Stephansson https://orcid.org/0000-0003-1528-4563

REFERENCES

1. Royal College of Obstetricians and Gynaecologists. Green-top Guideline No. 45. Birth After Previous Caesarean Birth. London: RCOG. 2015.
2. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins – Obstetrics. ACOG practice bulletin no. 205: vaginal birth after cesarean delivery. *Obstet Gynecol*. 2019;133:e110-e127.
3. Landon MB, Hauth JC, Leveno KJ, et al. Maternal and perinatal outcomes associated with a trial of labor after prior cesarean delivery. *N Engl J Med*. 2004;351:2581-2589.
4. Cunningham FG, Bangdiwala SI, Brown SS, et al. NIH consensus development conference draft statement on vaginal birth after cesarean: new insights. *NIH Consens State Sci Statements*. 2010;27:1-42.
5. Lindblad Wollmann C, Ahlberg M, Saltvedt S, Johansson K, Elvander C, Stephansson O. Risk of repeat cesarean delivery in women undergoing trial of labor: a population-based cohort study. *Acta Obstet Gynecol Scand*. 2018;97:1524-1529.
6. Landon MB, Leindecker S, Spong CY, et al. The MFMU Cesarean Registry: factors affecting the success of trial of labor after previous cesarean delivery. *Am J Obstet Gynecol*. 2005;193:1016-1023.
7. Lewkowitz AK, Nakagawa S, Thiet MP, Rosenstein MG. Effect of stage of initial labor dystocia on vaginal birth after cesarean success. *Am J Obstet Gynecol*. 2015;213(6):861.e1-861.e5.
8. Chung A, Macario A, El-Sayed YY, Riley ET, Duncan B, Druzin ML. Cost-effectiveness of a trial of labor after previous cesarean. *Obstet Gynecol*. 2001;97:932-941.
9. Gilbert SA, Grobman WA, Landon MB, et al. Lifetime cost-effectiveness of trial of labor after cesarean in the United States. *Value Health*. 2013;16:953-964.
10. Grobman WA, Peaceman AM, Socol ML. Cost-effectiveness of elective cesarean delivery after one prior low transverse cesarean. *Obstet Gynecol*. 2000;95:745-751.
11. Grobman WA, Lai Y, Landon MB, et al. Prediction of uterine rupture associated with attempted vaginal birth after cesarean delivery. *Am J Obstet Gynecol*. 2008;199:30.e1-5.
12. Grobman WA, Lai Y, Landon MB, et al. Development of a nomogram for prediction of vaginal birth after cesarean delivery. *Obstet Gynecol*. 2007;109:806-812.
13. Grobman W, Lai Y, Landon M, et al. Does information available at admission for delivery improve prediction of vaginal birth after cesarean? *Am J Perinatol*. 2009;26:693-701.

14. Fagerberg MC, Marsal K, Kallen K. Predicting the chance of vaginal delivery after one cesarean section: validation and elaboration of a published prediction model. *Eur J Obstet Gynecol Reprod Biol*. 2015;188:88-94.
15. Caughey AB, Shipp TD, Repke JT, Zelop C, Cohen A, Lieherman E. Trial of labor after cesarean delivery: the effect of previous vaginal delivery. *Am J Obstet Gynecol*. 1998;179:938-941.
16. Deo RC. Machine learning in medicine. *Circulation*. 2015;132:1920-1930.
17. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216-1219.
18. Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. *JAMA*. 2016;315:551-552.
19. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med*. 2003;26:172-181.
20. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. New York, NY: Springer-Verlag; 2013.
21. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*. 2014;25:418-426.
22. Rose S. A machine learning framework for plan payment risk adjustment. *Health Serv Res*. 2016;51:2358-2374.
23. Kominiarek MA, VanVeldhuisen P, Gregory K, Fridman M, Kim H, Hibbard JU. Intrapartum cesarean delivery in nulliparas: risk factors compared by two analytical approaches. *J Perinatol*. 2015;35:167-172.
24. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.
25. Stephansson O, Sandstrom A, Petersson G, Wikstrom AK, Cnattingius S. Prolonged second stage of labour, maternal infectious disease, urinary retention and other complications in the early postpartum period. *BJOG*. 2016;123:608-616.
26. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. *J Comput Graph Stat*. 2006;15:651-674.
27. Strobl C, Boulesteix A-L, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008;9:307.
28. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B*. 1996;58(1):267-288.
29. Mays VM, Ponce NA, Washington DL, Cochran SD. Classification of race and ethnicity: implications for public health. *Ann Rev Public Health*. 2003;24:83-110.
30. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight – reconsidering the use of race correction in clinical algorithms. *N Engl J Med*. 2020;383:874-882.
31. Strobl C, Boulesteix A-L, Zeileis A, Hothorn T. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*. 2007;8:25.
32. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer Series in Statistics*. New York: Springer; 2009.
33. Wu E, Kaimal AJ, Houston K, Yee LM, Nakagawa S, Kuppermann M. Strength of preference for vaginal birth as a predictor of delivery mode among women who attempt a vaginal delivery. *Am J Obstet Gynecol*. 2014;210(5):440.e1-440.e6.
34. Yee LM, Liu LY, Grobman WA. Relationship between obstetricians' cognitive and affective traits and delivery outcomes among women with a prior cesarean. *Am J Obstet Gynecol*. 2015;210(5):440.e1-440.e6.
35. Thornton P. Limitations of vaginal birth after cesarean success prediction. *J Midwifery Womens Health*. 2018;63:115-120.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Lindblad Wollmann C, Hart KD, Liu C, Caughey AB, Stephansson O, Snowden JM. Predicting vaginal birth after previous cesarean: Using machine-learning models and a population-based cohort in Sweden. *Acta Obstet Gynecol Scand*. 2021;100:513–520. https://doi.org/10.1111/aogs.14020