

## OBSTETRICS

## Are prediction models for vaginal birth after cesarean accurate?



Benjamin S. Harris, MD, MPH; R. Phillips Heine, MD; Jinyoung Park, MA; Keturah R. Faurot, MPH, PhD; Maeve K. Hopkins, MD; Andrew J. Rivara, MD; Hanna R. Kemeny, BS; Chad A. Grotegut, MD, MSc, MBA; J. Eric Jelovsek, MD, MMed

**BACKGROUND:** The use of trial of labor after cesarean delivery calculators in the prediction of successful vaginal birth after cesarean delivery gives physicians an evidence-based tool to assist with patient counseling and risk stratification. Before deployment of prediction models for routine care at an institutional level, it is recommended to test their performance initially in the institution's target population. This allows the institution to understand not only the overall accuracy of the model for the intended population but also to comprehend where the accuracy of the model is most limited when predicting across the range of predictions (calibration).

**OBJECTIVE:** The purpose of this study was to compare 3 models that predict successful vaginal birth after cesarean delivery with the use of a single tertiary referral cohort before continuous model deployment in the electronic medical record.

**STUDY DESIGN:** All cesarean births for failed trial of labor after cesarean delivery and successful vaginal birth after cesarean delivery at an academic health system between May 2013 and March 2016 were reviewed. Women with a history of 1 previous cesarean birth who underwent a trial of labor with a term ( $\geq 37$  weeks gestation), cephalic, and singleton gestation were included. Women with antepartum intrauterine fetal death or fetal anomalies were excluded. The probability of successful vaginal birth after cesarean delivery was calculated with the use of 3 prediction models: Grobman 2007, Grobman 2009, and Metz 2013 and compared with actual vaginal birth after cesarean delivery success. Each model's performance was measured with the use of concordance indices, Brier scores, and calibration plots. Decision curve analysis identified the

range of threshold probabilities for which the best prediction model would be of clinical value.

**RESULTS:** Four hundred four women met the eligibility criteria. The observed rate of successful vaginal birth after cesarean delivery was 75% (305/404). Concordance indices were 0.717 (95% confidence interval, 0.659–0.778), 0.703 (95% confidence interval, 0.647–0.758), and 0.727 (95% confidence interval, 0.669–0.779), respectively. Brier scores were 0.172, 0.205, and 0.179, respectively. Calibration demonstrated that Grobman 2007 and Metz vaginal birth after cesarean delivery models were most accurate when predicted probabilities were  $>60\%$  and were beneficial for counseling women who did not desire to have vaginal birth after cesarean delivery but had a predicted success rates of 60–90%. The models underpredicted actual probabilities when predicting success at  $<60\%$ . The Grobman 2007 and Metz vaginal birth after cesarean delivery models provided greatest net benefit between threshold probabilities of 60–90% but did not provide a net benefit with lower predicted probabilities of success compared with a strategy of recommending vaginal birth after cesarean delivery for all women.

**CONCLUSION:** When 3 commonly used vaginal birth after cesarean delivery prediction models are compared in the same population, there are differences in performance that may affect an institution's choice of which model to use.

**Key words:** calculator, calibration, decision curve analysis, prediction model, risk, TOLAC, validation, VBAC model

The use of trial of labor after cesarean delivery (TOLAC) calculators in the prediction of successful vaginal birth after cesarean delivery (VBAC) gives physicians an evidence-based tool to assist with patient counseling and risk stratification. Such prediction models estimate the individualized probability of successful VBAC with the use of patient demographics, comorbidities, and obstetric/perinatal risk factors.<sup>1</sup> As a result, providers are able to give

individualized guidance regarding the chance of successful TOLAC while balancing the risk of maternal/fetal morbidity that is associated with failed TOLAC vs elective repeat cesarean delivery.

Before the deployment of prediction models for routine care at an institutional level, it is recommended initially to test their performance in the institution's target population.<sup>2</sup> This allows the institution to understand not only the overall accuracy of the model for the intended population but also to comprehend where the accuracy of the model is most limited when predicting across the range of predictions (calibration).

Several published models are available to predict the probability of a successful VBAC in women who undergo TOLAC;

3 models (Grobman 2007 model, Grobman 2009 model, and Metz 2013 model) are used commonly in clinical practice. We searched MEDLINE from inception to August 27, 2018, using a mix of keywords and medical subject headings for “vaginal birth after cesarean” in combination with terms that are related to prediction models, such as “prediction,” “validation,” “risk,” and “theoretical models.” We also searched for articles citing all 3 models using Scopus and Web of Science. Although we did identify 4 citations that cited all 3 models,<sup>3–6</sup> we were unable to identify studies that simultaneously have compared performance of all 3 models in the same cohort. Moreover, decision curve analysis was not reported for any of the models. Decision curve analysis

**Cite this article as:** Harris BS, Heine RP, Park J, et al. Are prediction models for vaginal birth after cesarean accurate? *Am J Obstet Gynecol* 2019;220:492.e1-7.

0002-9378/\$36.00

© 2019 Elsevier Inc. All rights reserved.

<https://doi.org/10.1016/j.ajog.2019.01.232>

## AJOG at a Glance

**Why was this study conducted?**

The purpose of this study was to compare 3 models that predict successful vaginal birth after cesarean delivery with the use of a single tertiary referral cohort before model deployment in the electronic medical record.

**Key findings**

Calibration demonstrated that the Grobman 2007 and Metz vaginal birth after cesarean delivery models were accurate when the predicted probabilities were >60% but were underpredicted actual probabilities when predicting success at <60%. Both models provided greatest net benefit between threshold probabilities of 60–90% but did not provide a net benefit with lower predicted probabilities of success compared with a strategy of all women attempting vaginal birth after cesarean delivery.

**What does this add to what is known?**

It is not beneficial to use these vaginal birth after cesarean delivery models to make threshold decisions in the lower success range (<60%). We suggest that other institutions study the models' calibrations in their population to choose the most accurate model for their patient population.

offers insight into clinical consequences of the use of each model at chosen threshold probabilities.

The primary aim of this study was to assess and compare the predictive performance of these 3 models with the use

of temporally and geographically distinct data from a single academic health system that were not used to develop any of the models. The secondary aim was to use decision curve analysis to inform clinicians about a range of threshold probabilities in which the models would be of clinical value once deployed in the health system.

**Materials and Methods**

All VBAC and cesarean deliveries for failed TOLAC that were performed at Duke University between May 2013 and March 2016 were identified through the electronic medical record perinatal database. Women with a history of 1 previous cesarean birth who underwent a trial of labor with a term ( $\geq 37$  weeks), cephalic, and singleton gestation were included. Women with antepartum intrauterine fetal death or fetal anomalies were excluded. Obstetric and medical data for each model's predictors and outcome were extracted from the electronic medical records by physicians. Predictors included maternal age, parity, gestational age at birth (by best obstetric estimate), self-reported race, marital status, and body mass index (kilograms per square meter) at admission for delivery. Body mass index was obtained from measurements that were performed at individual clinic visits and/or hospital admissions. In our population, the Grobman 2007, Grobman 2009, and Metz models used body mass index measurements from first prenatal visit, prepregnancy body mass index, and delivery body mass index (eg, most recent within 2 weeks of delivery), respectively. Additional obstetric predictors included hypertensive disorders of pregnancy (preeclampsia, eclampsia, gestational hypertension, chronic hypertension), gestational diabetes mellitus, use of oxytocin, estimated fetal weight (grams), cervical examination (dilation, effacement, station) during admission for delivery, induction of labor, indication for previous cesarean delivery, and history of previous vaginal birth. Women were considered to have a recurring indication for previous cesarean delivery if the indication was arrest of descent or dilation. If a woman had an unclear

**TABLE 1**  
**Characteristics of validation cohort**

Characteristic	Successful vaginal birth after cesarean delivery (n = 305)	Failed trial of labor after cesarean delivery (n = 99)
Maternal age, y <sup>a</sup>	30±5	31± 6
Missing, n (%)	0	0
Gestational age at delivery (wk) <sup>a</sup>	39.0±1.1	39.1±1.2
Missing, n (%)	2 (0.01)	0
Race, n (%)		
Hispanic	113 (0.37)	33 (0.33)
White	67 (0.22)	17 (0.17)
Black	87 (0.29)	35 (0.35)
Asian	21 (0.07)	7 (0.07)
Other	2 (0.01)	2 (0.02)
Unknown	15 (0.05)	5 (0.05)
Missing	0	0
Body mass index, n (%)		
<18.5 kg/m <sup>2</sup>	6 (0.02)	0
18.5–24.9 kg/m <sup>2</sup>	98 (0.38)	24 (0.32)
25–29.9 kg/m <sup>2</sup>	71 (0.28)	25 (0.33)
>30	81 (0.32)	30 (0.40)
Missing	49 (0.16)	24 (0.24)

Harris et al. External validation of VBAC prediction models. *Am J Obstet Gynecol* 2019.

(continued)

**TABLE 1**  
**Characteristics of validation cohort** (continued)

Characteristic	Successful vaginal birth after cesarean delivery (n = 305)	Failed trial of labor after cesarean delivery (n = 99)
Admission cervical examination		
Dilation <sup>a</sup>	4±3	3±3
Missing, n (%)	2 (0.01)	6 (0.06)
Effacement <sup>a</sup>	69±24	61±28
Missing, n (%)	3 (0.01)	6 (0.06)
Station <sup>a</sup>	4±2	4±2
Missing, n (%)	6 (0.02)	6 (0.06)
History of vaginal delivery, n (%)	164 (0.54)	27 (0.28)
Missing	2 (0.01)	1 (0.01)
Previous vaginal birth after cesarean birth, n (%)	106 (0.35)	12 (0.12)
Missing	0	0
Recurrent indication for primary cesarean delivery, n (%)	86 (0.29)	51 (0.52)
Missing	35 (0.11)	10 (0.10)
Induction of labor, n (%)	79 (0.26)	39 (0.40)
Missing	2 (0.01)	11 (0.11)
Maternal preeclampsia or eclampsia, n (%)	19 (0.06)	12 (0.12)
Missing	3 (0.01)	4 (0.04)
Maternal gestational diabetes mellitus, n (%)	19 (0.06)	6 (0.07)
Missing	2 (0.01)	2 (0.02)
Neonatal birthweight, g <sup>a</sup>	3305±509	3341±503
Missing, n (%)	5 (0.02)	22 (0.22)

<sup>a</sup> Data are given as mean±standard deviation.

Harris et al. External validation of VBAC prediction models. Am J Obstet Gynecol 2019.

indication for previous cesarean delivery, it was labeled as missing. Bishop score was calculated based on admission cervical examination. The outcome was a

successful VBAC (dichotomous: yes/no). The failed TOLAC cohort included women who attempted a trial of labor and underwent cesarean delivery for

arrest disorders, failed induction, or nonreassuring fetal heart tracing.

Predictors of women who underwent a successful VBAC were compared with those with failed TOLAC. Missing predictor values were imputed with multivariate imputation by chained equations.<sup>7,8</sup> Predictive mean matching was used for numeric data; logistic regression was used for binary data, and polytomous (unordered) regression was used for factors >2 levels.

The probability of a successful VBAC was calculated with each model's published algorithm. The Grobman prediction models use information that is available at the woman's first prenatal visit and admission for delivery.<sup>9,10</sup> Similarly, each woman's data at admission for delivery was used to calculate a VBAC score with the Metz model.<sup>1</sup> Each score was mapped to the probability of successful VBAC using Table 3 from the original publication.<sup>1</sup>

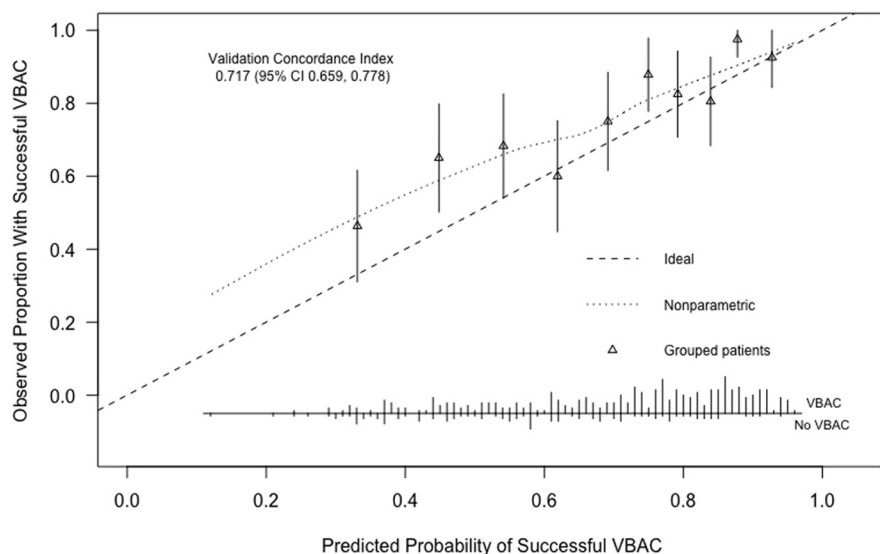
Model performance was measured by 3 methods: (1) a concordance index, (2) a Brier score, and (3) a calibration plot. The concordance index is determined by ranking an individual's risk among all subjects in the cohort (discrimination). It measures the model's ability to assign a higher predicted probability to a patient who is at high risk and a lower predicted probability to a patient who is at low risk. The concordance index (or c-statistic) is equivalent to the area under the receiver operating characteristic curve for a dichotomous outcome. The Brier score is similar to the  $R^2$  value because it measures overall accuracy of a model's predictions and is useful when combined with calibration.<sup>11</sup> Calibration refers to how closely the predicted probability of

**TABLE 2**  
**Comparison of model performance in the validation cohort**

Model	N	Published concordance index (95% confidence interval)	Validation concordance index (95% confidence interval)	Validation Brier score
Grobman et al, 2007 <sup>9</sup>	7666	0.754 (0.742–0.755)	0.717 (0.659–0.778)	0.172
Grobman et al, 2009 <sup>10</sup>	9616	0.774 (0.764–0.784)	0.703 (0.647–0.758)	0.205
Metz et al, 2013 <sup>1</sup>	1170	0.80 (0.76–0.84)	0.727 (0.669–0.779)	0.179

Harris et al. External validation of VBAC prediction models. Am J Obstet Gynecol 2019.

**FIGURE 1**  
Calibration plot of the Grobman 2007 model

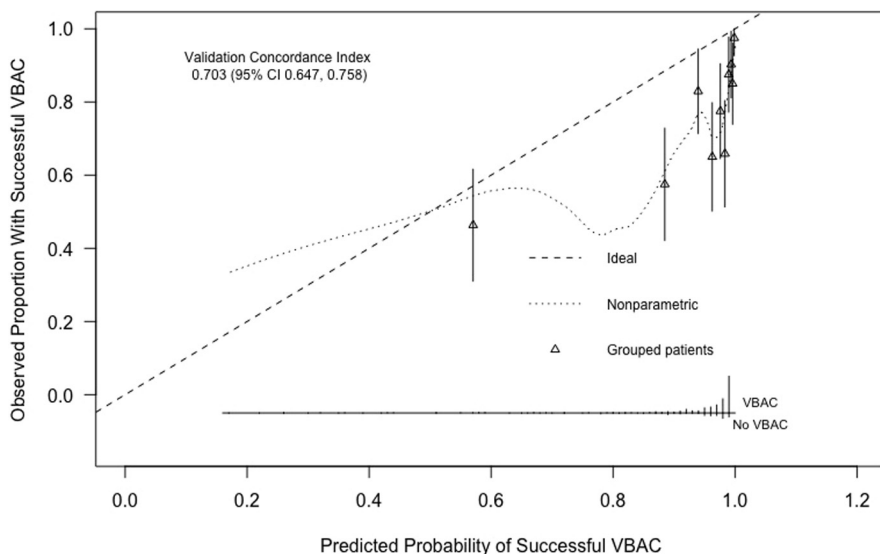


Model predictions were less accurate at lower predicted probabilities (0.1–0.4 and 0.5–0.7) compared with higher probabilities (>0.7). The rug plot allows for visually quantitating the relative number of persons with and without the outcome across the range of predictions.

VBAC, vaginal birth after cesarean delivery.

Harris et al. External validation of VBAC prediction models. *Am J Obstet Gynecol* 2019.

**FIGURE 2**  
Calibration plot of Grobman 2009 model



There was poor calibration of the model in the study cohort. Most of the predictions were >80% predicted probability with few predictions below this threshold. Additionally, predicted probabilities from 0.6–1.0 were lower than actual proportions of successful vaginal birth after cesarean delivery. The rug plot allows for visually quantitating the relative number of persons with and without the outcome across the range of predictions.

VBAC, vaginal birth after cesarean delivery.

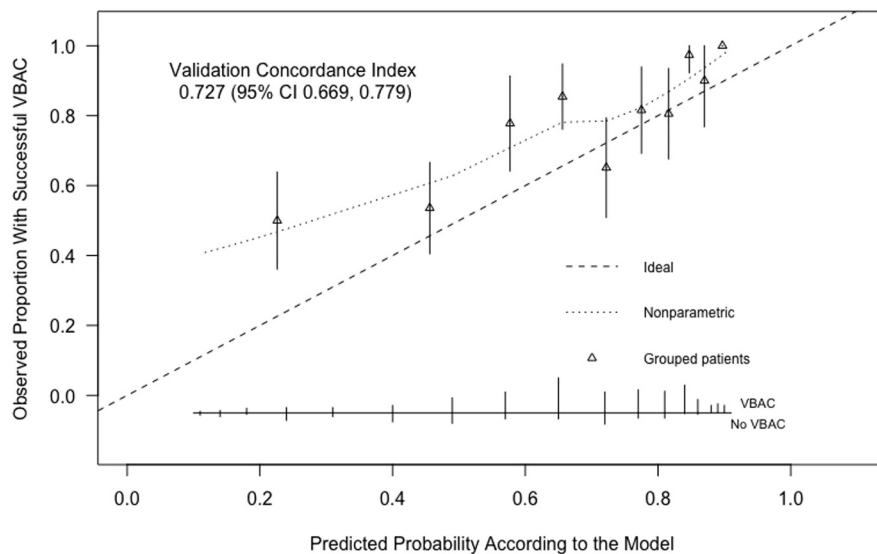
Harris et al. External validation of VBAC prediction models. *Am J Obstet Gynecol* 2019.

an outcome agrees with the observed probability of an outcome. Perfect model calibration is represented graphically by a slope of 1 and an intercept of 0.<sup>2</sup> Calibration curves were prepared that plot the predicted probability of successful VBAC against the actual probability of successful VBAC. This allowed for the determination of whether each model's predicted probability was too high or low (calibration) compared with actual probability at a range of predicted probabilities. Proportions within each decile were also identified on each plot.

A decision curve was plotted to inform clinicians about the range of threshold probabilities for which the prediction model would be of clinical value.<sup>12</sup> Decision curve analysis offers insight into clinical consequences of using the model by determining the relationship between a chosen predicted probability threshold and the relative value of false-positive and false-negative results to obtain a value of net benefit of the use of the model at that threshold.<sup>13</sup> This is useful before institutional deployment because clinicians are likely to have different predicted probability thresholds for recommending clinical decisions. All analyses were performed using R (version 3.4.3) and guidelines of transparent reporting of a multivariable prediction model for individual prognosis or diagnosis were followed.<sup>2</sup> The study was approved by the Institutional Review Board.

## Results

Four hundred four women met eligibility criteria. The observed rate of successful VBAC was 75% (305/404). Women in the successful VBAC cohort were predominantly Hispanic (37.1%) and had normal body mass indexes (38.3%). Fifty-four percent of the participants with successful VBAC had a history of vaginal delivery (Table 1). Twenty-nine percent of those with successful VBAC had recurrent indication for cesarean delivery compared with 52% among patients with failed TOLAC ( $P<.001$ ). Among those women who were admitted for a TOLAC, the rate of successful VBAC was higher among women who were in labor compared

**FIGURE 3**  
**Calibration plot of Metz model**

Most of the predicted probabilities were lower than the actual proportions in women with successful vaginal birth after cesarean delivery. The predictions were less accurate at lower predicted probabilities (0.1–0.4 and 0.5–0.7) compared with higher probabilities (>0.7). The rug plot allows for visually quantitating the relative number of persons with and without the outcome across the range of predictions.

VBAC, vaginal birth after cesarean delivery.

Harris et al. External validation of VBAC prediction models. *Am J Obstet Gynecol* 2019.

with those who were induced (225/285 [79%] vs 79/119 [66%];  $P=.006$ ).

The Grobman 2007 model validation concordance index was 0.717 (95% confidence interval [CI], 0.659–0.778), and Brier score was 0.172.<sup>9</sup> The discriminatory ability was similar to the published concordance index of 0.754 (95% CI, 0.742–0.755; Table 2). The calibration curve is depicted in Figure 1. The predicted probabilities were lower than the actual proportions of women with successful VBAC among all predictions. The predictions were less accurate at lower predicted probabilities (0.2–0.6) compared with higher probabilities (>0.6; Figure 1). For example, at a predicted probability of successful VBAC of 50%, the observed VBAC frequency was 60% in our population.

The Grobman 2009 model sample concordance index was 0.703 (95% CI, 0.647–0.758), and Brier score was 0.205.<sup>10</sup> This discriminatory ability was lower than the published concordance index of 0.774 (95% CI, 0.764–0.784; Table 2). The calibration curve is

depicted in Figure 2. There was poor calibration of the model in the study cohort. The majority of predictions were >80% predicted probability, with few predictions below this threshold. Additionally, predicted probabilities of 0.6–1.0) were lower than actual proportions of successful VBAC (Figure 2).

The Metz 2013 model sample concordance index was 0.727 (95% CI, 0.669–0.779), and Brier score was 0.179.<sup>1</sup> This discriminatory ability was similar to the published concordance index of 0.800 (95% CI, 0.760–0.840; Table 2). The calibration curve is depicted in Figure 3. Performance was similar to the Grobman 2007 model. The majority of predicted probabilities were lower than the actual proportions in women with successful VBAC. The predictions were less accurate at lower predicted probabilities (0.1–0.4 and 0.5–0.7) compared with higher probabilities (>0.7)

Figure 4 allows comparison of the 3 decision curves for each model. There are 5 curves: (1) allow everyone to

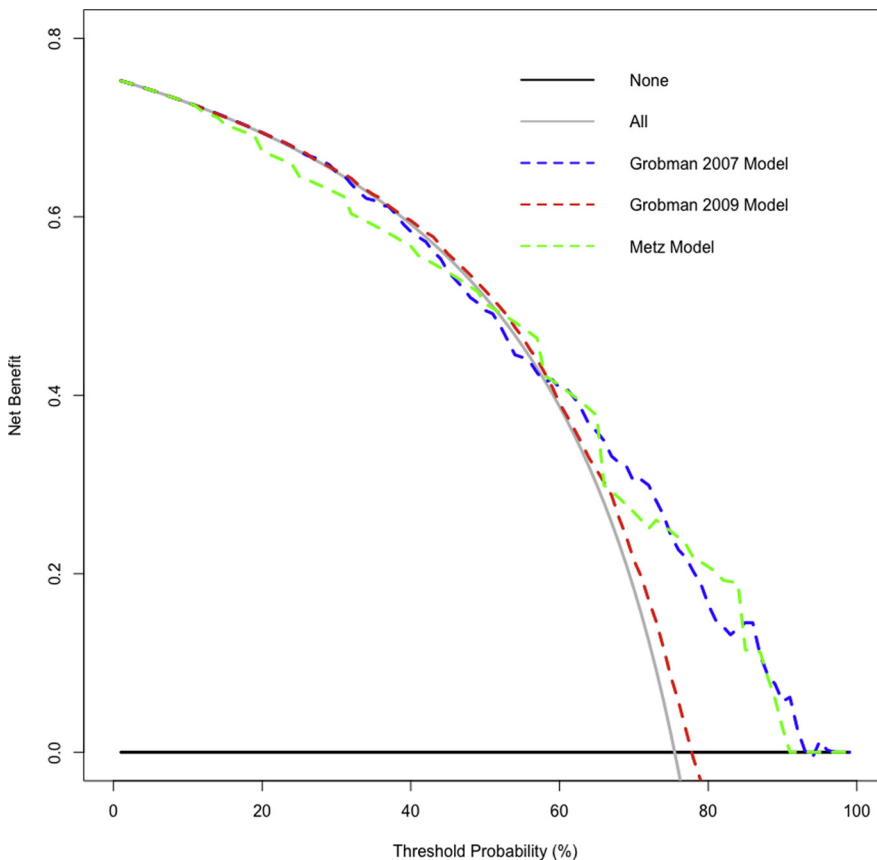
attempt VBAC (All), (2) allow no one to attempt VBAC (None), or (3) allow women who meet a range of threshold probabilities with the use of the Grobman 2007 model, (4) Grobman 2009 model, or (5) Metz model. The Grobman 2007 model has a higher net benefit at 65–80% threshold probabilities, which means that it is beneficial to use the model to make decisions in patients who fall within this range of threshold probabilities. The Metz model has a higher net benefit at <65% and from 80–90% threshold probabilities. For the purposes of counseling, the Metz model is not beneficial when the predicted probability of success is <65%.

## Comment

### Main findings

This study demonstrated the value of understanding and comparing the performances of 3 commonly used VBAC models in an institutional cohort before routine deployment. Although these findings are specific to a single health system, we recommend other institutions undertake similar institution-specific studies before widely applying any prediction model for routine care. These results led our institution to conclude the following: (1) We can reassure clinicians and our institutional leadership that the Grobman 2007 and Metz VBAC models are accurate when probabilities are >60% and that probabilities in this range can be communicated confidently to patients and their care providers to inform decisions. This provides useful information to patients who do not desire to attempt VBAC but have predicted success rates of 60–90%. (2) None of the models were accurate when probabilities were <60%. For example, providing a predicted risk of successful VBAC of 20% may not be much different from a predicted probability of 50%. Therefore, it would not be beneficial and could potentially be harmful if these probabilities were used to make threshold decisions in this lower success range. We plan to modify the displayed result of the calculator to report that “this patient’s predicted probability is <60% and the model is not accurate in

**FIGURE 4**  
**Decision Curves of 3 models that predict successful vaginal birth after cesarean birth using a threshold of  $p_t$**



The Figure displays the net benefit curves for the 3 models. If the threshold of approximately 60% was used to designate an individual as high probability of successful vaginal birth after cesarean delivery after a trial of labor instead of allowing all patients to trial of labor, there will be no net benefit of the use of any of the 3 models compared with a strategy of not using any model. If the threshold of approximately 70% was used to designate an individual as high probability of successful vaginal birth after cesarean delivery instead of allowing all patients to a trial of labor, the Grobman 2007 model identified 5 more cases per 100 of unnecessary trial of labor after cesarean delivery. The reduction in unnecessary trial of labor after cesarean delivery per 100 patients is calculated in the following manner: (net benefit of the model—net benefit of allowing all to trials of labor after cesarean delivery)/ $p_t/[1-p_t] \times 100$ . This value is net of false negatives and is therefore equivalent to the reduction in unnecessary trials of labor after cesarean delivery without a decrease in the number of patients with successful vaginal birth after cesarean delivery who have a trial of labor after cesarean delivery.

$p_t$ , threshold probability; VBAC, vaginal birth after cesarean delivery.

Harris et al. External validation of VBAC prediction models. *Am J Obstet Gynecol* 2019.

this range.” We suggest other institutions study the model’s calibration in their population to understand and potentially avoid areas of miscalibration.<sup>3</sup> The Grobman 2009 model had unacceptable performance in our population, and we do recommend its use at our institution. Similar studies at other institutions may benefit users of these models.

### Implications of findings

Model performance at our institution represented the natural history of model application in our population. That is, patients were not counseled routinely regarding TOLAC with the use of any of the published prediction models. At Duke University during the study time period, the average number of annual deliveries was 3250. Our counseling practice is to

allow all women with 1 previous low transverse cesarean delivery to try a trial of labor if they desire, while discussing the risks of failed TOLAC compared with elective repeat cesarean delivery. If a prediction model is used in counseling by house staff, advanced practice providers, or maternal fetal medicine specialists, the Grobman 2007 model is used. Counseling is initiated routinely at the first prenatal visit by our obstetric providers.

The Grobman 2007 and Metz models had similar overall discriminatory ability (concordance indices,  $>0.7$ ) in our population compared with their respective original validation cohorts; however, calibration demonstrated that most of the model’s overall accuracy was attributed to predictions of  $>60\%$ , and both models underpredicted actual success when predictions were  $<60\%$ . Similar to our findings, other investigators demonstrated that the Grobman 2007 model accurately performs at higher predicted success rates,<sup>14,15</sup> but not at lower probabilities. For example, Maykin et al<sup>16</sup> performed an external validation in a population of 568 women at a single academic institution in California. This study demonstrated accurate model performance with predicted success rates  $>65\%$ . Among women with predicted success rates  $<35\%$ , the model underestimated actual VBAC success rates by nearly 2-fold, which suggests poor model accuracy at lower predicted probabilities.<sup>16</sup> It is possible that this variability in accuracy may be partly due to geographic and racial or ethnic diversity, but it may also be that the model is miscalibrated at lower predicted probabilities of success. Either way, it is important to understand how any model performs across the entire range of a specific population before routine use.

### Strengths and limitations

The strengths of our study are the comparison of commonly used models with the use of the same population and the use of decision curve analysis to aid in counseling patients as they approach decisions regarding TOLAC. Decision curve analysis is a useful method to compare prediction models when using the same validation cohort. It are also

useful in quantitating the overall net benefit to patients when decisions are made with the use of a range of predicted thresholds<sup>12</sup> because clinicians and patients at different institutions are likely to have different thresholds. The Grobman 2007 model and Metz model added net benefit when decisions are made using the model by a clinician or patient with predicted probabilities of 60–95%. Therefore, predictions from both models are most useful for the patient who is predicted to have a >60% chance of success but is unsure about attempting VBAC. In this case, the model could reassure the patient about success. However, use of these model predictions below threshold probability of 60% do not provide any net benefit over allowing all patients to attempt VBAC.

This study has limitations. The validation was retrospective by design, had a smaller cohort sample compared with the original validation studies, and relied on electronic medical record data entry by varying providers to best reflect routine clinical care. The data set does not include women who were eligible for a TOLAC but who declined and elected to proceed with an elective repeat cesarean delivery. Because of this, we are not able to calculate the proportion of eligible women at Duke University Health System who elected to attempt a trial of labor. Next, a power analysis was not conducted because, unfortunately, there are no generally accepted approaches to estimate the sample size requirements for validation studies of prediction models.<sup>2</sup> It is possible that poor performance of the Grobman 2009 model was due to differences in how cervical examination data were collected from the electronic health record in our cohort with respect to the original publication. Similar to previous studies, we are also not able to account for the effect of counseling, patient preference, or individual labor management styles on delivery outcomes.<sup>1</sup> Despite these

limitations, we believe that this process is a similar and a feasible way for institutions to validate models developed by other researchers before deployment at their own institution.

### Conclusion/future research direction

In conclusion, institutions should use similar methods to validate a prediction model's performance before routine use in their population. When multiple models exist, this process can be used to choose which model to deploy, establish baseline overall accuracy, and understand areas of model miscalibration in their patient population. We acknowledge that this requires an institution to have expertise in model recalibration. In our health system, the Grobman 2007 and Metz VBAC models were accurate when probabilities are >60%, potentially providing useful information to patients who do not desire to VBAC but have predicted success rates of 60–90%. The 2 models are not accurate when predicted success rates were <60%. Clinicians are encouraged to test their VBAC model's performance in the target population and encourage systematic routine timepoints for checking the model's performance over time. ■

### References

1. Metz TD, Stoddard GJ, Henry E, Jackson M, Holmgren C, Esplin S. Simple, validated vaginal birth after cesarean delivery prediction model for use at the time of admission. *Obstet Gynecol* 2013;122:571–8.
2. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multi-variable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
3. 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.
4. Dresang LT, Hampton A. Vaginal birth after cesarean (VBAC) calculator risks. *Birth* 2015;42:379–80.
5. Edgoose J, Hampton A, Schrager S. How good are online TOLAC calculators for

predicting successful VBAC? *Evid Based Pract* 2015;18:5–6.

6. Warren JB, Hamilton A, Ruge B. What's the best way to predict the success of a trial of labor after a previous C-section? *J Fam Pract* 2015;64:E3–7.
7. Van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;16:219–42.
8. Buuren SV, Groothuis-Oudshoorn K. MICE multivariate imputation by chained equations in R. *J Stat Soft* 2011;45:1–67.
9. Grobman WA, Lai Y, Landon MB, et al. Development of a nomogram for prediction of vaginal birth after cesarean delivery. *Obstet Gynecol* 2007;109:806–12.
10. Grobman WA, Lai Y, Landon MB, et al. Does information available at admission for delivery improve prediction of vaginal birth after cesarean? *Am J Perinatol* 2009;26:693–701.
11. Ruffin K. Use of Brier score to assess binary predictions. *J Clin Epidemiol* 2010;63:938–9.
12. Vickers AJ. Decision analysis for the evaluation of diagnostic tests, prediction models and molecular markers. *Am Stat* 2008;62:314–20.
13. Van Calster B, Vickers AJ. Calibration of risk prediction models: impact on decision-analytic performance. *Med Decis Making* 2015;35:162–9.
14. Metz TD, Allshouse AA, Faucett AM, Grobman WA. Validation of a vaginal birth after cesarean delivery prediction model in women with two prior cesarean deliveries. *Obstet Gynecol* 2015;125:948–52.
15. Chaillet N, Bujold E, Dubé E, Grobman WA. Validation of a prediction model for vaginal birth after caesarean. *J Obstet Gynaecol Can* 2013;35:119–24.
16. Maykin MM, Mularz AJ, Lee LK, Valderramos SG. Validation of a prediction model for vaginal birth after cesarean delivery reveals unexpected success in a diverse American population. *AJP Rep* 2017;7:e31–8.

### Author and article information

From the Department of Obstetrics and Gynecology, Duke University Health System (Drs Harris, Heine, Hopkins, Rivara, Grotegut, and Jelovsek) and Duke University School of Medicine (Ms Kemery), Durham, NC; the Department of Physical Medicine and Rehabilitation, University of North Carolina, Chapel Hill, NC (Dr Faurot and Ms Park).

Received Sept. 26, 2018; revised Jan. 5, 2019; accepted Jan. 28, 2019.

The authors report no conflict of interest.

Corresponding author: Benjamin S. Harris, MD, MPH. [benjamin.harris@duke.edu](mailto:benjamin.harris@duke.edu)