



Published in final edited form as:

*Obstet Gynecol.* 2011 April ; 117(4): 805–811. doi:10.1097/AOG.0b013e3182114ad2.

## Using a Simplified Bishop Score to Predict Vaginal Delivery

S. Katherine Laughon, MD, MS<sup>1</sup>, Jun Zhang, PhD, MD<sup>1</sup>, James Troendle, PhD<sup>1</sup>, Liping Sun, MD, MS<sup>1</sup>, and Uma M. Reddy, MD, MPH<sup>1</sup>

<sup>1</sup>Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD

### Abstract

**Objective**—Determination of the Bishop score is the most commonly used method to assess the readiness of the cervix for induction. However, it was created without modern statistical methods. Our objective was to determine whether a simplified score can predict vaginal delivery equally well.

**Methods**—Data were analyzed for 5,610 nulliparous women with singleton, uncomplicated pregnancies between 37 0/7 – 41 6/7 weeks undergoing labor induction. These women had all five components of the Bishop score recorded. Logistic regression was performed and a simplified score created with significant components. Positive and negative predictive values (PPV and NPV) and positive likelihood ratio (LR+) were calculated.

**Results**—In the regression model, only dilation, station and effacement were significantly associated with vaginal delivery ( $P < .01$ ). The simplified Bishop score was then devised using these 3 components (range 0 – 9) and compared to the original Bishop score (range 0 – 13) for prediction of successful induction, resulting in vaginal delivery. Compared to the original Bishop score  $> 8$ , the simplified Bishop score  $> 5$  had a similar or better PPV (87.7% versus 87.0%), NPV (31.3% versus 29.8%), LR+ (2.34 versus 2.12) and correct classification rate (51.0% versus 47.3%). Application of the simplified Bishop score in other populations including indicated induction and spontaneous labor at term and preterm were associated with similar vaginal delivery rates compared to the original Bishop score.

**Conclusion**—The simplified Bishop score comprised of dilation, station and effacement attains a similarly high predictive ability of successful induction as the original score.

### INTRODUCTION

In the 1960s Dr. Edward Bishop developed a pelvic scoring system using cervical dilatation, effacement, station, consistency and position with a possible range from 0–13.<sup>1</sup> Based on clinical experience, he concluded that elective induction in multiparous women with uncomplicated pregnancies at term was successful with a score of  $> 8$ . Shortly after the Bishop score was introduced, other investigators created weighting for the components of the score, and found that cervical dilation was more associated with the time of latent phase compared to the other components. However, the weighted Bishop score did not provide a clinically significant improvement in predicting duration of labor compared to the original score.<sup>2, 3</sup> New scores have been proposed, the Bishop score has been modified, and attempts have been made to improve the Bishop score by adjusting for additional maternal and

---

Corresponding author: Dr. S. Katherine Laughon, Epidemiology Branch, NICHD, National Institutes of Health, 6100 Executive Blvd., Room 7B03, Rockville, MD 20852, Tel: 301-435-6935, laughonsk@mail.nih.gov.

This research has been accepted as a poster presentation at the Annual Meeting of the Society for Maternal-Fetal Medicine, San Francisco, CA, February 10, 2011.

obstetrical characteristics, but these scores in general have not proven to be superior to the original score, and these more cumbersome scores have not been widely adapted into busy clinical practice.<sup>4-8</sup> The Bishop score remains the most commonly used system to assess for pre-induction readiness.<sup>9</sup>

Since the original Bishop score was created on an empiric basis without modern statistical methods and the five components are correlated, the question remains whether all components are necessary in predicting vaginal delivery. If only some of the components are independently associated with successful induction, then the score can be reduced to contain only those components with equivalent ability to predict a successful induction. Our objective was to determine whether a simplified Bishop score can predict vaginal delivery equally well in nulliparous women with uncomplicated pregnancies undergoing induction of labor at term in contemporary obstetrical practice. We then investigated whether a simplified Bishop score could be applied for other indications for induction and at different gestational ages.

## MATERIAL AND METHODS

The Consortium on Safe Labor was a study conducted by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health involving 228,668 deliveries between 2002 and 2008 from 12 clinical centers and 19 hospitals.<sup>10</sup> Institutional Review Board approval was obtained by all participating institutions. Data were collected from electronic medical records including demographics, past medical history, labor and delivery information as well as obstetrical, post partum and neonatal outcomes. Additional data from the neonatal intensive care unit were collected and linked to the newborn record. The patient data were supplemented with maternal and newborn discharge ICD-9 codes for each delivery. Each site transferred data in electronic format to the data coordinating center where data were mapped to common categories for each pre-defined variable. Data were cleaned and logic checking performed. Validation studies indicated that the electronic medical records were an accurate representation of the medical charts.<sup>10</sup>

Eleven sites provided indications for induction. We included nulliparous women with a singleton gestation, delivering between 37 0/7 – 41 6/7 weeks of gestation, with vertex presentation, and were uncomplicated pregnancies undergoing elective or postdates induction of labor, or induction for precursors that could have been expectantly managed, including uncomplicated gestational hypertension<sup>11</sup> or chronic hypertension prior to 39 weeks of gestation<sup>12</sup>, history of maternal, obstetrical or fetal indication in a prior pregnancy or induction for suspected fetal macrosomia without diabetes<sup>13</sup>. We excluded women with a previous uterine scar (n = 12), stillbirth (n = 16), any infant with congenital anomalies (n = 795) or who had an induction for any other reason including chorioamnionitis, fetal compromise, maternal preeclampsia, maternal medical conditions, and vaginal bleeding. A total of 12,996 women were available for final data analysis and of these, 5610 women had all five components of the Bishop score and this was designated the “training” population.

Logistic regression with backwards elimination was performed to investigate which components of the Bishop score (dilation, effacement, station, consistency and position) were significantly associated with successful vaginal delivery in a model adjusted for site. A simplified Bishop score was created by comparing the regression coefficients and using only the components that had a final P< .01 by Wald test. The significance level of P<.01 for an effect to stay in the model was chosen because while P<.05 might be statistically significant, the purpose of the study was to simplify the score. We chose to include only those components that were the main contributors to success of vaginal delivery. The regression

model for the simplified Bishop score was validated using a bootstrap method with samples of the same size as the original dataset.<sup>14</sup> Bootstrapping is a technique that allows a given population to be randomly resampled to create multiple datasets of the same size. The analysis was re-run in each bootstrap sample to evaluate whether our decision making regarding choice of which of the five components of the original Bishop score to include in a simplified score was robust. Logistic regression was performed with  $P < .01$  significance level for the effect to stay in the model in a backward elimination step using the dataset from each of the 1000 bootstrap samples.

Interactions were explored between the components that were statistically significant and Spearman correlation coefficients were calculated. Sensitivity, specificity, positive and negative predictive values (PPV and NPV) and likelihood ratio positive (LR+) were calculated for the original Bishop score and the simplified Bishop score. The correct classification rate was calculated by adding the number of true positives and true negatives/total number of subjects classified.

The simplified Bishop score was compared to the original Bishop score in two test populations where women had all cervical components present: at term (37 0/7 – 41 6/7 weeks' gestation) and preterm (32 6/7 – 36 6/7 weeks' gestation) undergoing an indicated induction of labor, including maternal, obstetrical or fetal indications for induction (for example, preeclampsia, maternal medical diseases, small for gestational age, oligohydramnios) and did not include any women in the training population. In order to test the Bishop score and simplified Bishop score in a "natural experiment", we also evaluated these scores in women with spontaneous labor at term (37 0/7 – 41 6/7 weeks' gestation) and preterm (32 0/7 – 36 6/7 weeks' gestation).

## RESULTS

There were 5,610 women included in the training population and their characteristics are presented in Table 1. Most women were between the ages of 18 and 34 years and had an average height between 60 and 68 inches. About 1/3 of women were overweight (BMI 25.0 to 29.9 kg/m<sup>2</sup>) at delivery and 38.9% of women were obese (BMI  $\geq$  30.0 kg/m<sup>2</sup>). The majority (69.3%) of women were white/non-Hispanic, followed by 10.9% black/non-Hispanic and 6.7% Hispanic. Induction of labor occurred more often in women with private insurance (77.0%), non-smokers (97.1%) and at or after 39 weeks of gestation. There were 1716 (30.6%) women who had a Bishop score  $> 8$  prior to induction.

### Creation of a Simplified Score

Overall, 75.3% women (n= 4,224) had a vaginal delivery. In the regression model, dilation had the highest regression coefficient (.45) followed by station (.32), and these cervical components were both highly significant ( $P < .001$  and  $P = .009$ , respectively, Table 2). Effacement had a regression coefficient that was similar to consistency (.15 versus .13, respectively), although effacement was highly significant ( $P < .001$ ) while consistency was not ( $P = .07$ ). Cervical position had a very small contribution to the model (regression coefficient = .01) and was not significant ( $P = .06$ ). There were no significant interactions between these components, although they were correlated (Spearman  $r = .3$  to  $.5$ ,  $P < .001$ ). We chose to include dilation, station and effacement in a simplified score since these were the cervical components that had the largest three regression coefficients and were highly significantly associated with success of vaginal delivery.

In order to validate the process of developing a simplified score, a bootstrap method was used. The bootstrap method resulted in dilation and station always being chosen in the

model, and effacement chosen for 70.5% of the different bootstrap samples, overall supporting our choice of cervical components from the regression model (Table 3).

### Comparison of the Simplified Score to the Original Bishop Score

At a given sensitivity and specificity for vaginal delivery, the PPV, NPV and correct classifications rates were similar to the original Bishop score compared to using a simplified Bishop score based on dilation, effacement and station only (Table 4). For example, using the original Bishop score  $> 8$ , the simplified Bishop score with the closest sensitivity and specificity would be  $> 5$ . Compared to the original Bishop score  $> 8$ , the simplified Bishop score  $> 5$  had a similar PPV (87.7% for the simplified versus 87.0% for the original score) and NPV (31.3% for the simplified versus 29.8% for the original score). The likelihood ratio positive test and the correct classification rate were also similar or slightly better (2.3% versus 2.2% and 51.0% versus 47.3%, respectively).

### Application in Different Populations

We then compared the simplified Bishop score to the original Bishop score for the following separate populations of women: term (37 0/7 – 41 6/7 weeks' gestation) indicated induction and spontaneous labor, and preterm (32 0/7 – 36 6/7 weeks' gestation) indicated induction and spontaneous labor. The simplified Bishop score was associated with a similar vaginal delivery rate compared to the original Bishop score (Figure). For illustration, a simplified Bishop score  $> 5$  performed similarly to an original Bishop score  $> 8$  in both the indicated inductions and spontaneous labor at term and preterm with the similar correct classification rates (Table 5).

## DISCUSSION

In nulliparous women with uncomplicated pregnancies undergoing an induction of labor at term, a simplified Bishop score with three components: dilation, station and effacement predicted vaginal delivery similarly to the original Bishop score. The simplified Bishop score also was comparable to the original Bishop score in predicting successful vaginal delivery in women with an indicated induction both at term and preterm between 32 – 36 6/7 weeks of gestation. Even in women who presented in spontaneous labor at term and preterm, the simplified Bishop score was similar to the original Bishop score, suggesting that the simplified score is equivalent to the original score in the setting that it was developed.

Other attempts at modifying or evaluating the Bishop score have used different outcomes such as length of labor or achieving active labor, and many included multiparous women who are known to have more successful inductions.<sup>3-5,7</sup> We chose vaginal delivery as the primary outcome, because this is what clinicians and patients define as success. Our study also has the advantage of having a large number of nulliparous women. Thus, we were able to use modern statistical methods to find which components of the Bishop score were independently associated with vaginal delivery in order to create a simplified score.

There is a possibility that women who had all five components of the Bishop score recorded are different in baseline characteristics from women who were missing some of the components. However most of the women (72.2%) had dilation, station, and effacement present, and many clinicians informally already use a simplified Bishop score. It is more likely that the recording of some versus five components of the Bishop score was based on clinician preference rather than something inherently different about a woman undergoing an induction. Given the large numbers we were able to test the simplified score in other populations of women, including indicated induction and spontaneous labor both term and preterm, and the simplified Bishop score performed similarly to the original Bishop score in

predicting vaginal delivery in all of these settings which suggests that missing cervical components were likely not an issue.

Our findings are similar to a prospective study of 134 women undergoing an induction of labor at term, where only the cervical components of dilation and effacement were associated with vaginal delivery within 24 hours.<sup>15</sup> Using an “abbreviated” Bishop score including dilation and effacement only  $> 3$ , the predictive characteristics of vaginal delivery (excluding 23 women who had an emergency cesarean delivery for maternal or fetal indications) were PPV 85.5%, NPV 65.7%, and LR + 2.61, which were similar to our simplified Bishop  $> 5$ . An older, smaller study of 40 nulliparous and 69 multiparous women also found that only dilation was associated with the length of latent phase of labor after labor induction.<sup>16</sup> Our study found both effacement and station to be significant in addition to dilation likely because we had a large number of women and thus more power. While the addition of position or consistency may be significantly associated with successful vaginal delivery in a different population of women, the purpose of our model was to simplify the score, so we chose only the components that were both highly significant in the regression and contributed the most to vaginal delivery as determined by the regression coefficients. Of note, simplifying the score even further by using only the two components with the highest regression coefficients, dilation and station, resulted in a worse correct classification rate compared to the simplified Bishop score using all three components of dilation, station and effacement (data not shown). Our findings are also supported by a secondary analysis of four randomized controlled trials with a total of 781 women comparing different induction methods for indicated induction after 37 weeks’ gestation, and the cervical components dilation, effacement and station were independently associated with vaginal delivery within 24 hours after adjusting for maternal and obstetrical characteristics, although only position and station were associated with spontaneous vaginal delivery.<sup>17</sup>

Other studies have created variations of the Bishop score. In a prospective study of 1189 women undergoing induction mostly for indicated indications, Lange et al. used linear regression to create a new score with the cervical components of dilation and station from the original Bishop score and length measured as centimeters as opposed to percentage, with dilation multiplied by two.<sup>7</sup> The indications for induction (PROM, amniotomy and medically induced) and definitions of failure (delivery within 24 hours or labor established within 8 hours for the medically induced group) were different from our study as well as a lower overall rate of failure of around 15% compared to 25% in our study. Nonetheless, Lange’s score was found to perform similarly to the original Bishop score in that population of women. Dhall et al. also created a new score in 200 women undergoing indicated induction with a slightly lower vaginal delivery rate (71.5%) than our study.<sup>8</sup> Dilation, effacement and consistency were rescored and weighted, and parity was also included. The Dhall score had higher prediction of success rate at both ends of the score, but the study was limited because no women had a Bishop score  $> 8$ . In addition, using a reasonably accurate prediction of Dhall score  $\geq 7$  which corresponded to a Bishop score cut-off point of 4, the Dhall score only performed significantly better in multiparous but not nulliparous women.

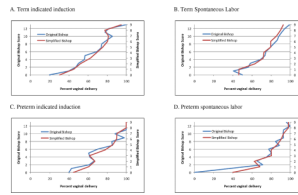
In summary, reassessing the original Bishop score using modern statistical methods resulted in a simplified score with only three components: dilation, station and effacement yielding an equivalently high predictive ability. The simplified Bishop score performed similarly to the original Bishop score in predicting vaginal delivery in indicated inductions term and preterm, as well as in spontaneous labor at term and preterm. Given that our study is a large, nationally representative cohort reflecting current clinical practice, our findings are generalizable. As cervical position and consistency do not add to the overall ability to predict vaginal delivery, we believe that the original Bishop score can be replaced with a simplified score using dilation, station and effacement only.

## Acknowledgments

The data included in this paper were obtained from the Consortium on Safe Labor, which was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, through Contract No. HHSN267200603425C. Institutions involved in the Consortium include, in alphabetical order: Baystate Medical Center, Springfield, MA; Cedars-Sinai Medical Center Burnes Allen Research Center, Los Angeles, CA; Christiana Care Health System, Newark, DE; Georgetown University Hospital, MedStar Health, Washington, DC; Indiana University Clarian Health, Indianapolis, IN; Intermountain Healthcare and the University of Utah, Salt Lake City, Utah; Maimonides Medical Center, Brooklyn, NY; MetroHealth Medical Center, Cleveland, OH.; Summa Health System, Akron City Hospital, Akron, OH; The EMMES Corporation, Rockville MD (Data Coordinating Center); University of Illinois at Chicago, Chicago, IL; University of Miami, Miami, FL; and University of Texas Health Science Center at Houston, Houston, Texas.

## References

1. Bishop EH. Pelvic Scoring for Elective Induction. *Obstet Gynecol.* 1964; 24:266–8. [PubMed: 14199536]
2. Friedman EA, Niswander KR, Bayonet-Rivera NP, Sachtleben MR. Relation of prelabor evaluation to inducibility and the course of labor. *Obstet Gynecol.* 1966; 28:495–501. [PubMed: 5925035]
3. Friedman EA, Niswander KR, Bayonet-Rivera NP, Sachtleben MR. Prelabor status evaluation. II. Weighted score. *Obstet Gynecol.* 1967; 29:539–44. [PubMed: 6021030]
4. Burnett JE Jr. Preinduction scoring: an objective approach to induction of labor. *Obstet Gynecol.* 1966; 28:479–83. [PubMed: 5925032]
5. Fields H. Induction of labor. Readiness for induction. *Am J Obstet Gynecol.* 1966; 95:426–9. [PubMed: 5939264]
6. Hughey MJ, McElin TW, Bird CC. An evaluation of preinduction scoring systems. *Obstet Gynecol.* 1976; 48:635–41. [PubMed: 995333]
7. Lange AP, Secher NJ, Westergaard JG, Skovgard I. Prelabor evaluation of inducibility. *Obstet Gynecol.* 1982; 60:137–47. [PubMed: 7155472]
8. Dhall K, Mittal SC, Kumar A. Evaluation of preinduction scoring systems. *Aust N Z J Obstet Gynaecol.* 1987; 27:309–11. [PubMed: 3453668]
9. Baaacke KA, Edwards RK. Preinduction cervical assessment. *Clin Obstet Gynecol.* 2006; 49:564–72. [PubMed: 16885663]
10. Zhang J, Troendle J, Reddy UM, Laughon SK, Branch DW, Burkman R, et al. Contemporary cesarean delivery practice in the United States. *Am J Obstet Gynecol.* 203:326 e1–26 e10. [PubMed: 20708166]
11. Sibai BM. Diagnosis and management of gestational hypertension and preeclampsia. *Obstet Gynecol.* 2003; 102:181–92. [PubMed: 12850627]
12. Sibai BM. Chronic hypertension in pregnancy. *Obstet Gynecol.* 2002; 100:369–77. [PubMed: 12151166]
13. Irion O, Boulvain M. Induction of labour for suspected fetal macrosomia. *Cochrane Database Syst Rev.* 2000:CD000938. [PubMed: 10796221]
14. Izrael, D.; Battaglia, AA.; Hoaglin, DC.; Battaglia, MP. SAS® Macros and Tools for Working with Weighted Logistic Regression Models That Use Survey Data. Seattle SAS Users Group International Proceedings; Seattle, Washington: 2003. <http://www2.sas.com/proceedings/sugi28/275-28.pdf>
15. Reis FM, Gervasi MT, Florio P, Bracalente G, Fadalti M, Severi FM, et al. Prediction of successful induction of labor at term: role of clinical history, digital examination, ultrasound assessment of the cervix, and fetal fibronectin assay. *Am J Obstet Gynecol.* 2003; 189:1361–7. [PubMed: 14634569]
16. Watson WJ, Stevens D, Welter S, Day D. Factors predicting successful labor induction. *Obstet Gynecol.* 1996; 88:990–2. [PubMed: 8942840]
17. Crane JM, Delaney T, Butt KD, Bennett KA, Hutchens D, Young DC. Predictors of successful labor induction with oral or vaginal misoprostol. *J Matern Fetal Neonatal Med.* 2004; 15:319–23. [PubMed: 15280123]



**Figure 1.**

Comparison of the original Bishop score (range 0 – 13, left Y-axis) to the simplified Bishop score with dilation, station and effacement only (range 0 – 9, right Y-axis) by vaginal delivery rate (X-axis) for the following separate populations of women: a. term (37 0/7 – 41 6/7 weeks of gestation) indicated induction; b. term spontaneous labor; c. preterm (32 0/7 – 36 6/7 weeks of gestation) indicated induction; and d. preterm spontaneous labor.

**Table 1**

Maternal and obstetrical characteristics of women in the training cohort.

Characteristics	Training Cohort n = 5,610 n (%)
Maternal age – year	
< 18	264 (4.7)
18–34	5,011 (89.3)
≥ 35	330 (5.9)
Missing	5 (0.1)
Body mass index at delivery - kg/m <sup>2</sup>	
< 25.0	537 (9.6)
25.0 – 29.9	1,858 (33.1)
≥ 30.0	2,183 (38.9)
Missing	1,032 (18.4)
Height – inches	
< 60	265 (4.7)
60 – 64	1,968 (35.1)
65 – 68	2,098 (37.4)
> 68	418 (7.5)
Missing	861 (15.3)
Race or ethnic group	
White/non-Hispanic	3,890 (69.3)
Black/non-Hispanic	611 (10.9)
Hispanic	378 (6.7)
Asian/Pacific Islander	193 (3.5)
Other/Unknown	538 (9.6)
Insurance type	
Private	4,319 (77.0)
Public	1,222 (21.8)
Self pay	21 (0.4)
Other/Unknown	48 (0.8)
Smoker	163 (2.9)
Membranes ruptured	994 (17.7)
Gestational age at delivery – weeks	
37	246 (4.4)
38	538 (9.6)
39	1,441 (25.7)
40	1,734 (30.9)
41	1,651 (29.4)
Dilation	
0	437 (7.8)
1 – 2	3,468 (61.8)
3 – 4	1,572 (28.0)



Characteristics	Training Cohort n = 5,610 n (%)
> 4	133 (2.4)
Effacement	
0 – 30	500 (8.9)
40 – 50	1,065 (19.0)
60 – 70	1,708 (30.4)
> 70	2,337 (41.7)
Station	
–5 to –3	1,383 (24.7)
–2	2,107 (37.5)
–1 or 0	1,998 (35.6)
1+	122 (2.2)
Consistency	
Firm	341 (6.1)
Medium	1,735 (30.9)
Soft	3,534 (63.0)
Cervical position	
Posterior	2,207 (39.3)
Mid	2,371 (42.3)
Anterior	1,032 (18.4)

**Table 2**

The original Bishop score with logistic regression results for outcome of vaginal delivery.

Component	Regression coefficient*	P value	Subscore			
			0	1	2	3
Dilation (cm)	.45	< .001	0	1-2	3-4	5-6
Station	.32	.009	-3	-2	-1 or 0	+1 or +2
Effacement (%)	.15	< .001	0-30	40-50	60-70	80
Consistency	.13	.07	Firm	Medium	Soft	
Position	.01	.06	Posterior	Mid	Anterior	

\*The regression coefficients and P-values are from the full model with all 5 components and adjusted for site. After elimination of consistency with backwards elimination, position was still not significant (P=.07).

**Table 3**

Bootstrap method results evaluating regression model with outcome of vaginal delivery.

Number of variables removed	Variables removed	Percent of samples
0	None of variables were removed (full model contains six predictors: dilation, effacement, station, consistency, position and site)	5.5
1	Consistency	17.3
	Position	10.7
	Effacement	6.0
2	Consistency + Position	37.0
	Position + Effacement	14.6
	Consistency + Effacement	2.9
3	Consistency + Position + Effacement	6.0

Bootstrap samples (n = 1,000) of the same size as the original dataset were used. Logistic regression was performed with P<.01 significance level of the Wald chi-square for an effect to stay in the model in a backward elimination step using the resample.

**Table 4**

Predictive characteristics of Bishop score versus simplified Bishop score.

Score	Sensitivity %	Specificity %	PPV %	NPV %	LR +	Correct classification rate %
Original Bishop						
> 5	75.1	49.2	81.8	39.3	1.47	68.7
> 6	62.9	62.8	83.8	35.7	1.69	62.9
> 7	49.8	75.8	86.2	33.1	2.05	56.2
> 8	35.3	83.8	87.0	29.8	2.19	47.3
Simplified Bishop						
> 3	75.8	49.3	82.0	40.1	1.49	69.4
> 4	59.2	67.9	84.9	35.3	1.84	61.4
> 5	40.6	82.6	87.7	31.3	2.34	51.0
> 6	18.8	94.2	90.8	27.6	3.25	37.4

PPV, positive predictive value; NPV, negative predictive value; LR+, likelihood ratio positive (calculated by sensitivity/(1 – specificity)); the correct classification rate was calculated by number of true positives + number of true negatives/number of subjects classified.

Original Bishop includes dilation, effacement, station, consistency and position with a possible range from 0 – 13; Simplified Bishop includes dilation, station and effacement only with a possible range from 0 – 9.

**Table 5**  
Application of simplified Bishop score in different populations for success of vaginal delivery.

Population	n*	Vaginal Delivery %	Sensitivity %	Specificity %	PPV %	NPV %	LR +	Correct classification rate %
Indicated induction at term	4106	65.5						
Original Bishop > 8			38.4	81.6	89.7	24.1	2.09	46.7
Simplified Bishop >5			39.2	81.2	87.5	28.3	2.08	48.8
Term spontaneous labor	18615	85.5						
Original Bishop > 8			22.5	89.4	51.1	70.2	2.13	67.4
Simplified Bishop >5			25.4	89.4	45.8	77.2	2.39	72.7
Indicated induction preterm	658	70.1						
Original Bishop > 8			34.4	95.0	97.5	20.6	6.9	43.6
Simplified Bishop >5			33.9	91.4	95.4	20.6	3.92	43.0
Preterm spontaneous labor	1334	88.8						
Original Bishop > 8			18.9	94.1	68.5	63.0	3.19	63.6
Simplified Bishop >5			19.4	93.6	63.8	66.7	3.03	66.3

PPV, positive predictive value; NPV, negative predictive value; LR+, likelihood ratio positive (calculated by sensitivity/(1 - specificity)); the correct classification rate was calculated by number of true positives + number of true negatives/number of subjects classified.

Original Bishop includes dilatation, effacement, station, consistency and position with a possible range from 0 - 13; Simplified Bishop includes dilation, station and effacement only with a possible range from 0 - 9.

Indicated induction was considered as all maternal, obstetrical or fetal indications for induction and did not include any women in the training population (elective, premature rupture of the membranes or for soft indication).

\* Number of women with all five components of the original Bishop score for that population.

Term includes delivery between 37 0/7 - 41 6/7; preterm includes delivery between 32 0/7 - 36 6/7.