Published in final edited form as:

Obstet Gynecol. 2008 September ; 112(3): 508–515. doi:10.1097/AOG.0b013e3181842087.

Preterm Prediction Study: Comparison of the Cervical Score and Bishop Score for Prediction of Spontaneous Preterm Delivery

RB Newman, RL Goldenberg, JD Iams, PJ Meis, BM Mercer, AH Moawad, E Thom, M Miodovnik, SN Caritis, M Dombrowski, and GR Thurnau for the National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units Network (MFMU)
From the Departments of Obstetrics and Gynecology at the University of Alabama, Birmingham, AL (RLG); Wake Forest University, Winston-Salem, North Carolina (PJM); University of Chicago, Chicago, IL (AHM); University of Cincinnati, Cincinnati, OH (MM); University of Pittsburgh, Pittsburgh, PA (SNC); Ohio State University, Columbus, OH (JDI); University of Oklahoma, Oklahoma City, OK (GRT); Medical University of South Carolina, Charleston, SC (RBN); University of Tennessee, Memphis, TN (BMM); Wayne State University, Detroit, MI (MD); and The George Washington University Biostatistics Center, Washington, DC (ET).

Abstract

OBJECTIVE—To prospectively compare digital Cervical Score with Bishop Score as a predictor of spontaneous preterm delivery before 35 weeks of gestation.

METHODS—Data from a cohort of 2,916 singleton pregnancies enrolled in a multicenter preterm prediction study were available. Patients underwent digital cervical examinations at 22–24 and 26–29 weeks of gestation for calculation of Bishop Score and Cervical Score. Relationships between Bishop Score, Cervical Score, and spontaneous preterm delivery were assessed with multivariable logistic regression analysis, McNemar's test, and receiver operating characteristics (ROC) curves to identify appropriate diagnostic thresholds and predictive capability.

RESULTS—One hundred twenty-seven of 2,916 patients (4.4%) undergoing cervical examination at 22–24 weeks had a spontaneous preterm delivery before 35 weeks. Eighty-four of the 2,538 (3.3%) re-examined at 26–29 weeks also had spontaneous preterm delivery. ROC curves indicated that optimal diagnostic thresholds for Bishop Score were at least 4 at 22–24 weeks and at least 5 at 26–29 weeks and less than 1.5 at both examinations for Cervical Score. At 22–24 weeks, areas under the ROC curve favored Bishop Score. At 26–29 weeks, there was no significant difference in areas under the ROC curve; however, a Cervical Score less than 1.5 (sensitivity 35.7%, false positive rate 4.8%) was superior to a Bishop Score ≥ 5 ($p < 0.001$) [DS1].

CONCLUSIONS—Both cervical evaluations are associated with spontaneous preterm delivery in a singleton population, however, predictive capabilities for spontaneous preterm delivery were modest among women with low event prevalence. Although Bishop Score performed better in the mid-trimester, by 26–29 weeks a Cervical Score less than 1.5 was a better predictor of spontaneous preterm delivery before 35 weeks than a Bishop Score at least 5.

Corresponding author: Roger B. Newman, M.D., Professor and Vice Chairman, Department of Obstetrics and Gynecology, Medical University of South Carolina, 96 Jonathan Lucas Street - MSC 619, Charleston, SC 29425, newmanr@musc.edu, FAX: 843-792-0533, OFFICE: 843-792-7100.

For members of the other members of the NICHD MFMU, see the Appendix online at www.greenjournal.org/cgi/content/full/112/3/xxx/DC1.

Financial Disclosure: The authors have no potential conflicts of interest to disclose.

INTRODUCTION

Prematurity is the major contributor to perinatal morbidity and mortality in the United States. Therefore, identifying women at risk for preterm delivery remains an issue of paramount importance. Recent investigators have described an inverse relationship between cervical length by endovaginal sonography and the likelihood of subsequent preterm birth (1–4). Digital cervical examination may also provide significant clinical information, while avoiding the cost and logistical difficulties of serial transvaginal sonography. We sought to estimate the best use of the information obtained by digital examination in terms of predicting preterm delivery risk.

Antepartum digital cervical examination traditionally has been performed using the Bishop Score calculated by assessment of dilatation, effacement, consistency of the cervix, its position, and the station of the presenting part (5). The Cervical Score described by Houlton in 1982 attempted to refine the information available from the digital cervical examination by replacing effacement with length as a descriptor of the unlabored cervix and ignoring the more subjective parameters of consistency, position and station (6). The Cervical Score places a greater emphasis on cervical length while cervical effacement is only one of five components of the Bishop Score which was originally developed as an evaluation of the inductibility of the cervix at term rather than as a predictor of preterm birth.

The objective of this investigation was to compare the ability of these two digital cervical assessment scores to predict spontaneous preterm delivery (SPTD) in a large low-risk singleton population. It is hypothesized that the Cervical Score would have superior predictive capability for SPTD.

MATERIALS AND METHODS

A prospective cohort study was conducted among 2,916 women with singleton pregnancies enrolled in a multi-center (10 sites) preterm prediction study between October 1992 and July 1994 by the National Institutes of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units Network. A secondary analysis was performed of data available from that study. The investigation was approved by the human subjects review board at each institution. All participating women provided written informed consent. Inclusion criteria were singleton gestations with intact membranes enrolled between 22 and 24 weeks' gestation. All women underwent an ultrasound examination before enrollment to confirm the last menstrual period dating criteria or to establish the duration of gestation if those criteria could not be confirmed. Race and parity distributions reflected each participating center with no single center contributing more than 20% of the total study population. Exclusion criteria were those pregnancies with fetal demise, congenital malformation, placenta previa, multifetal gestation, cervical cerclage, HIV positivity, preterm labor, preterm premature rupture of the membranes, prolapse of the fetal membranes, or plans to deliver away from the clinical center. The sample size was calculated assuming a premature delivery rate of 3.5%; that at least 5% of the women would have positive results on any given screening test, and that the odds ratio for premature delivery was 2.0 or more for women with positive results on that screening test as compared with women with negative results. A sample of 3000 women were chosen to give a lower 95% CI limit greater than 1 for this odds ratio.

All 2,916 enrolled patients underwent a second trimester digital cervical examination at 22–24 weeks' gestation. Of these, 2,538 underwent repeat digital cervical examination in the early third trimester between 26–29 weeks. The digital exams were performed by designated study personnel at each clinical site. The study personnel were trained and certified according to study protocol. Prior to initiation of the study each center designated one examiner to be the “standard” to which all cervical examiners were compared. This person was usually the

institutional principal investigator but it had to be an individual with at least 5 years experience in cervical examinations. For a period of time the “standard” examiner and study nurse assessed cervical dilation, length, station, consistency and position together to establish consistent agreement between both examiners. At that point, ten patients were examined independently and the detailed results recorded on the Cervical Examination Standardization Form by each examiner. These forms were collected and mailed to the George Washington Biostatistics Coordinating Center for certification (GWBCC). The examinations were compared for consistency between the examiners using pre-specified criteria for agreement. If the results of the paired exams were not consistent, the study nurse was required to complete more cervical examinations in conjunction with the “standard” examiner. During the study, approximately every 12 weeks, a sample of patients were randomly chosen by the GWBCC, for verification of continuing consistency as for the initial certification. All study personnel conducting cervical examinations had to be both initially and continuously certified. A manual of operations was developed describing the technique for evaluating each element of the Bishop Score. The recorded findings from each cervical examination were used for calculation of both the Bishop Score and Cervical Score centrally by the GWBCC (Table 1) (5,6).

The relationship between the Bishop Score, Cervical Score, and SPTD <37, <35, and <32 weeks’ gestation was assessed using multivariable logistic regression analysis adjusting for body mass index, black race and previous preterm delivery to estimate odds ratios and 95% confidence intervals (95% CI) at both assessments (22–24 and 26–29 weeks). Further adjustment for clinical center did not change the results. SPTD <35 weeks’ gestation was selected as the primary outcome due to its relative frequency and the increased risk of neonatal morbidity with delivery earlier than this gestational age. SPTD was defined as those deliveries following spontaneous onset of preterm labor or preterm premature rupture of membranes.

Receiver Operating Characteristics (ROC) curves were created for both Bishop Score and Cervical Score at both assessments. Appropriate cut-points or diagnostic thresholds for both scores at each assessment were identified by visual inspection. Sensitivity, specificity, false positive rate (i.e. 100%-specificity) and positive and negative predictive values for tests using these thresholds were estimated. A comparison of the ability of each test to classify patients correctly according to whether they experienced SPTD < 35 weeks’ gestation or not was assessed using McNemar’s test for the chosen cut-points and in comparison to transvaginal sonographic cervical assessments previously performed.¹ To examine the overall performance of Bishop Score and Cervical Score for the entire range of cut-points, the areas under the ROC curves for Bishop Score and Cervical Score and the 95% confidence intervals were calculated at each assessment and compared (7). For all tests, a nominal 2-tailed p-value of < 0.05 was considered significant. No adjustments were made for multiple comparisons.

RESULTS

One hundred twenty-seven of the 2,916 enrolled patients (4.4%) undergoing digital cervical examination at 22–24 weeks had a SPTD <35 weeks’ gestation. The demographic characteristics of this entire cohort have been previously described (1,3). Baseline characteristics for those women who did (N=127) and did not (N=2789) experience SPTD <35 weeks’ gestation are compared in Table 2.

Of the original 2916 patients, 2538 were re-examined at 26–29 weeks. Of these, 84 had a SPTD <35 weeks’ gestation. The Bishop Score and the Cervical Score were both significantly associated with SPTD using an adjusted multivariable logistic regression analysis (Table 3). An increased odds of spontaneous preterm delivery at <37, <35, and <32 weeks’ gestation was observed per unit increase in the Bishop Score at both the 22–24 week examination and the 26–29 week re-examination (Table 3). A decreased odds of spontaneous preterm delivery at

<37, <35, and <32 weeks' gestation was observed per unit increase in the Cervical Score at both examinations (Table 3). These odds ratios are not directly comparable since the units for the two scores are not comparable.

As the cut-point for the Bishop Score increases and as the cut-point for the Cervical Score decreases, the specificity, positive predictive values and odds ratios for SPTD <35 weeks' gestation increase, however the sensitivity quickly declines for both tests. It is noted that a very small proportion (< 10%) of the 127 patients with a SPTD < 35 weeks had a Bishop Score ≥ 6 or a Cervical Score < 1.0 between 22 and 24 weeks' gestation (Table 4). At 26–29 weeks' gestation, similar results were observed (Table 5). The areas under the ROC curves ranged between 0.61 and 0.68 (random assignment would have an area of 0.5). However, all adjusted odds ratios for SPTD <35 weeks' gestation exceeded 2.9 for abnormal examinations defined by Bishop Score between 4 and 7 for 22–24 weeks and 5 – 7 for 26–29 weeks, or Cervical Score between 0 and 1.5 for both 22–24 and 26–29 weeks' gestation.

Inspection of the ROC curve inflection points suggested that the optimal diagnostic thresholds for SPTD <35 weeks at the 22–24 week examination were ≥ 4 for the Bishop Score and <1.5 for the Cervical Score. The ROC curves for the 22–24 week examination are shown in Figure 1. At 22–24 weeks' gestation, the area under the ROC curve was significantly larger for the Bishop Score than for the Cervical Score (0.66 versus 0.61, $p=0.03$), indicating that overall, Bishop Score is a better diagnostic test at this gestation. However, McNemar's test revealed that a Cervical Score < 1.5 was superior to a Bishop Score ≥ 4 ($p<0.001$) for the prediction of SPTD < 35 weeks' gestation. For 86.6% of the patients, the timing of delivery was correctly predicted by both tests and for 4.4% of the patients both tests were wrong. For 7.4% of the patients the Cervical Score was correct while the Bishop Score was wrong, and for 1.6% of the patients the Bishop Score was correct while the Cervical Score was wrong.

At the 26–29 week re-examination, the chosen diagnostic thresholds for SPTD <35 weeks were ≥ 5 for the Bishop Score and < 1.5 for the Cervical Score. The ROC curves for the 26–29 week examination are shown in Figure 2. The area under the ROC curves were not statistically significantly different for the Bishop Score and the Cervical Score (both 0.68, $p=0.90$), however, McNemar's test again revealed that a Cervical Score < 1.5 was superior to a Bishop Score ≥ 5 at 26–29 weeks' gestation for prediction of SPTD < 35 weeks ($p < 0.0001$) for prediction of SPTD <35 weeks'. For 89.0 % of the patients, the timing of delivery was correctly predicted by both tests and for 4.8% of the patients both tests were wrong. For 4.3% of the patients the Cervical Score was correct while the Bishop Score was wrong, and for 2.0% of the patients the Bishop Score was correct while the Cervical Score was wrong.

Sonographic cervical length ≤ 20 mm, funneling of the endocervical canal, and Cervical Score < 1.5 at 26 – 29 weeks' gestation were compared as predictive tests for SPTD < 35 weeks (Table 6). McNemar's test revealed superiority at 26–29 weeks for predicting SPTD < 35 weeks for a Cervical Score < 1.5 compared with whether or not funneling was present on endovaginal ultrasound ($p < 0.0001$). McNemar's test revealed no test superiority when comparing a Cervical Score < 1.5 and a cervical length ≤ 20 mm measured by ultrasound. However, when the Cervical Score < 1.5 and the cervical length ≤ 20 mm are compared in Table 6, it is noted that given the same specificity the cervical score had a slightly higher sensitivity.

DISCUSSION

Although digital cervical examination has been proposed by some as a routine method of assessing preterm delivery risk (8–10), it has not been generally accepted for this role in the United States. Many consider the components of digital cervical examination to be unacceptably subjective with high intra- and inter-observer variability (11–12). Additionally,

digital evaluation is limited to only the vaginal portion of the uterine cervix. With the development of endovaginal cervical sonography, it has been documented that digital examination underestimates the true cervical length by 10–20 mm (12–14). The supravaginal portion of the cervix immediately adjacent to the internal cervical os appears to be the part of the cervix that undergoes the earliest prelabor cervical change (1–3,15). Regardless of the clinician's experience, it is possible that effacement of the cervix beginning in the region of the internal cervical os may not be detectable by digital palpation until significant change has already occurred.

While it is true the supravaginal portion of the cervix and the internal cervical os can be assessed more directly by the ultrasound probe than by the examining finger, it does not necessarily follow that digital examination is therefore an inferior predictor of preterm delivery risk. Measures of test function for a digital examination obtained at 26–29 weeks' gestation are very similar to those obtained from the same patients with endovaginal sonographic cervical length measurements ≤ 20 mm or the presence of endocervical funneling (Table 6). It has been previously published that endovaginal sonographic cervical length measurements correlate significantly ($p < 0.001$ by the Jonckheere – Terpstra test) with the Bishop Score.¹ Subjects with a Bishop Score of 6 or more at 24 or 28 weeks' gestation had mean cervical lengths of 23.1 ± 13.3 mm and 25.4 ± 11.2 mm respectively.¹ Differences in predictive capability that exist between endovaginal sonography and digital cervical examination are small and may not justify the expense of endovaginal sonography in low-risk obstetrical populations especially after 24 weeks' gestation. As a result, it becomes important to identify the most effective use of the information obtained by digital examination. In this investigation, while Bishop Score was an overall better test than Cervical Score at 22–24 weeks, a Cervical Score of < 1.5 in the early third trimester (26–29 weeks) was a better test for SPTD < 35 weeks' gestation compared to a Bishop Score ≥ 5 .

The superiority of the Cervical Score cutoff at 26–29 weeks is most likely a function of its focus on cervical length and dilatation as opposed to the Bishop Score, which includes somewhat more subjective parameters such as cervical consistency, position, and station of the presenting part. It is important to remember that the Bishop Score was developed to estimate the inducibility of the cervix at term and not as a predictor of preterm delivery risk (5). When the components of the Bishop Score are evaluated individually, dilatation and effacement have been found to be the two best predictors of both inducibility and gestational age at delivery (16,17). The Cervical Score on the other hand was developed specifically to assess preterm delivery risk (6) and has been evaluated primarily in high-risk gestations. In multiple gestations, it was demonstrated that as the Cervical Score decreases, the mean time until delivery shortens (18,19). A Cervical Score ≤ 0 between 20–28 weeks' gestation predicted a 92% risk of preterm delivery (19). It must be cautioned, however, that test functions for either the Cervical Score or the Bishop Score will be significantly different in the multifetal versus low-risk singleton populations due to their much higher prevalence of premature delivery.

A possible limitation of this investigation does relate to the risk status of this population. Given the demographics of the participating centers the cohort is of relatively poor socioeconomic status with a high percentage of African-American women. Almost 16% also had a history of prior SPTD. This cohort may be at higher baseline risk than a private practice, community based population. These population differences would be expected to increase the test's positive predictive value. Another limitation may be that the predictive values of both the Cervical Score and the Bishop Score were estimated based on routine scheduled examinations.³ The clinical use of the digital cervical examination will usually be in symptomatic women. The anticipated greater risk of women undergoing indicated clinical examination may alter the predictive qualities of this test.

In this singleton population, neither Cervical Score nor Bishop Score can be considered a strong diagnostic test. It is conventional wisdom that to be a useful test, the area under the ROC curve should equal or exceed 0.75 – this level was not achieved by either score at either assessment. If the observed results held true, then among one thousand patients, a Cervical Score <1.5 compared with a Bishop Score ≥ 5 at 26–29 weeks would have correctly predicted a spontaneous preterm delivery in only one additional patient, although the Cervical Score would have correctly provided a reassuring result for about 20 additional women. The primary advantage of the Cervical Score is its relatively superior specificity which allows for greater reassurance when the examination is normal. This difference, as well as the overall validity testing, is not compelling enough to justify routine digital examination of low risk patients to determine a Cervical Score. However, cervical examinations are frequently performed in response to subjective maternal signs and symptoms. When a digital examination is performed for whatever reason, its predictive capability should be maximized. The Cervical Score provides the most information from that examination and, as an added benefit, is also an easier way to quantify cervical status.

A Cervical Score < 1.5 identified between 26–29 weeks' gestation would identify a very small portion of this population ($<6\%$) as being at risk for early preterm delivery. One out of five of these women would deliver prior to 35 weeks' gestation. Cervical Scores that are even lower are a rare finding in the early third trimester but are associated with significantly higher positive predictive values. Although, we are limited in our ability to prevent preterm labor, identification of women “at risk” is a prerequisite to any intervention. Identification of at-risk women early in the third trimester may allow the opportunity to employ interventions such as increased surveillance for preterm labor and consideration for administration of antenatal corticosteroids. A low Cervical Score may also identify an “at risk” group of women who might benefit from weekly injections of 17 alpha-hydroxy progesterone caproate (20). Consideration should be given to using a Cervical Score <1.5 in the mid or early third trimester as an inclusion criteria to study 17 alpha-hydroxy progesterone caproate and/or other interventions designed to reduce the risk of preterm birth.

Acknowledgments

Supported by grants from the National Institute of Child Health and Human Development (HD21410, HD21414, HD27860, HD27861, HD27869, HD27883, HD27889, HD27905, HD27915, HD27917, HD19897 and HD36801).

The authors thank Madeline Rice, PhD, for statistical analysis, Anita Das, PhD for data management and statistical analysis, and Rachel Copper RN and Frances Johnson, RN, for protocol development and coordination between clinical research centers.

REFERENCES

1. Iams JD, Goldenberg RL, Meis PJ, Mercer BM, Moawad A, Das A, et al. The length of the cervix and the risk of spontaneous premature delivery. *New Engl J Med* 1996;334:567–572. [PubMed: 8569824]
2. Anderson HF, Nugent CE, Wanty FD, Hayashi RH. Prediction of risk for preterm delivery by ultrasonographic measurement of the cervical length. *Am J Obstet Gynecol* 1990;163:859–867. [PubMed: 2206073]
3. Goldenberg RL, Iams JD, Mercer BM, Meis PJ, Moawad AH, et al. The Preterm Prediction Study: The value of new vs standard risk factors in predicting early and all spontaneous preterm births. *Am J Public Health* 1998;88:233–238. [PubMed: 9491013]
4. Imseis HM, Albert TA, Iams JD. Identifying twin gestations at low risk for preterm birth with a transvaginal ultrasonographic cervical measurement at 24–26 weeks' gestation. *Am J Obstet Gynecol* 1997;177:1149–1155. [PubMed: 9396911]
5. Bishop EH. Pelvic scoring for elective induction. *Obstet Gynecol* 1964;24:266–268. [PubMed: 14199536]

6. Houlton MCC, Marivate M, Philpott RH. Factors associated with preterm labour and changes in the cervix before labour in twin pregnancy. *Br J Obstet Gynaecol* 1982;89:190–194. [PubMed: 7066256]
7. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 1988;44:837–845. [PubMed: 3203132]
8. Wood C, Bannerman RHO, Booth RT, Pinkerton JHM. The prediction of premature labor by observation of the cervix and external tocography. *Am J Obstet Gynecol* 1965;91:396–402. [PubMed: 14259927]
9. Papiernik E, Bouyer J, Collin D, Winisdoerffer G, Dreyfuss J. Precocious cervical ripening and preterm labor. *Obstet Gynecol* 1986;67:238–242. [PubMed: 3945433]
10. Stubbs TM, VanDorsten JP, Miller MC. The preterm cervix and preterm labor: Relative risks, predictive values, and change over time. *Am J Obstet Gynecol* 1986;155:829–834. [PubMed: 3766637]
11. Holcomb WL Jr, Smeltzer JS. Cervical effacement: Variation and belief among clinicians. *Obstet Gynecol* 1991;78:43–45. [PubMed: 2047066]
12. Jackson GM, Ludmir J, Bader TJ. The accuracy of digital examination and ultrasound in the evaluation of cervical length. *Obstet Gynecol* 1992;79:214–218. [PubMed: 1731287]
13. Sonek JD, Iams JD, Blumenfeld M, Johnson F, Landon M, Gabbe S. Measurement of cervical length in pregnancy: Comparison between vaginal ultrasonography and digital examination. *Obstet Gynecol* 1990;76:172–175. [PubMed: 2196494]
14. Goldberg J, Newman RB, Rust PF. Interobserver reliability of digital and endovaginal ultrasonographic cervical length measurements. *Am J Obstet Gynecol* 1997;177:853–858. [PubMed: 9369833]
15. Iams JD, Johnson FF, Sonek J, Sachs L, Gebauer C, Samuels P. Cervical competence as a continuum: A study of cervical length and obstetric performance. *Am J Obstet Gynecol* 1995;172:1097–1106. [PubMed: 7726247]
16. Bouyer J, Papiernik E, Dreyfuss J, Collin D, Winisdoerffer B, Gueguen S. Maturation signs of the cervix and the prediction of preterm birth. *Obstet Gynecol* 1986;68:209–214. [PubMed: 3737036]
17. Lange JP, Secher NJ, Westergaard JG, Skovgard I. Prelabor evaluation of inducibility. *Obstet Gynecol* 1982;60:137–146. [PubMed: 7155472]
18. Neilson JP, Verkuy DAA, Crowther CA, Bannerman C. Preterm labor in twin pregnancies: Prediction by cervical assessment. *Obstet Gynecol* 1988;72:719–723. [PubMed: 3173923]
19. Newman RB, Godsey RK, Ellings JM, Campbell BA, Eller DP, Miller MC. Quantification of cervical change: Relationship to preterm delivery in the multifetal gestation. *Am J Obstet Gynecol* 1991;165:264–267. [PubMed: 1872326]
20. Meis PJ, Klebanoff M, Thom E, Dombrowski MP, Sibai B, Moawad AH, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxy progesterone caproate. *N Eng J Med* 2003;348:2379–2385.

Figure 1.

Receiver operating characteristics curve for spontaneous preterm birth (spontaneous preterm delivery) before 35 weeks of gestation using increasing Bishop Score (open circles) compared with a decreasing Cervical Score (open diamonds) obtained by digital cervical examination at 22–24 weeks' gestation. BS, Bishop score; CS, Cervical score.

Figure 2.

Receiver operating characteristics curve for spontaneous preterm birth (spontaneous preterm delivery) before 35 weeks of gestation for an increasing Bishop Score (open circles) compared with a decreasing Cervical Score (open diamonds) obtained by digital cervical examination at 26–29 weeks of gestation. CS, Cervical score; BS, Bishop Score

Table 1

Numeric Calculation of both the Bishop Score and a Cervical Score

Score	0	1	2	3
Dilation (cms)	Closed	1–2	3–4	≥ 5
Effacement %	0–30	40–50	60–70	≥ 80
Station *	–3	–2	–1, 0	+1, +2
Cervical Consistency	Firm	Medium	Soft	
Position of Cervix	Posterior	Midposition	Anterior	

Bishop Score = Dilation (cm) + Effacement (%) + Station + Cervical consistency + Position of cervix. Cervical Score = Cervical length (cms) – Cervical dilation (cms) at the internal os (i.e., a cervix which is 2 cm long but 1 cm dilated at the internal os is a cervical score =1).

* Based on a –3 to +3 scale

Table 2

Baseline characteristics of the study cohort by spontaneous preterm delivery < 35 weeks' gestation.

Baseline Characteristics	Patients with Spontaneous Preterm Birth <35 Weeks' Gestation (n=127)	Patients without Spontaneous Preterm Birth <35 Week's Gestation (n=2789)
Maternal age, years, mean $\pm$ standard deviation	23.6 $\pm$ 5.3	23.7 $\pm$ 5.5
Body mass index, kg/m ² , mean $\pm$ standard deviation	26.4 $\pm$ 5.9	28.7 $\pm$ 7.0
Education less than 12 years, number (percent)	57 (44.9%)	969 (34.7%)
Government insurance, number (percent)	115 (90.6%)	2494 (89.4%)
Married, number (percent)	28 (22.1%)	790 (28.3%)
Race or ethnic groups, number (percent)		
Black	89 (70.1%)	1742 (62.5%)
Hispanic	0 (0.0%)	23 (0.8%)
Other, non-white	0 (0.0%)	32 (1.2%)
Nulliparous, number (percent)	40 (31.5%)	1168 (41.9%)
Previous preterm delivery, number (percent)	59 (46.5%)	399 (14.3%)
Cigarette smoking during pregnancy, number (percent)	40 (31.5%)	854 (30.6%)
Alcohol during pregnancy, number (percent)	11 (8.7%)	312 (11.2%)

Table 3

Adjusted* logistic regression per unit increase in the Bishop and Cervical Scores for the outcome of SPTD at < 32, < 35, and < 37 weeks' gestation.

Bishop Score			Cervical Score	
22–24 wks	p-value	OR (95% CI)	p-value	OR (95% CI)
SPTD <32 wks	<0.0001	1.59 (1.33–1.89)	0.0010	0.58 (0.42–0.80)
SPTD <35 wks	<0.0001	1.46 (1.29–1.65)	<0.0001	0.62 (0.50–0.78)
SPTD <37 wks	<0.0001	1.29 (1.18–1.41)	0.0006	0.76 (0.65–0.89)
26–29 wks				
SPTD <32 wks	0.0806	1.24 (0.97–1.58)	0.0007	0.52 (0.36–0.76)
SPTD <35 wks	<0.0001	1.43 (1.25–1.64)	<0.0001	0.47 (0.37–0.60)
SPTD <37 wks	<0.0001	1.28 (1.17–1.39)	<0.0001	0.64 (0.54–0.75)

* Adjusted for body mass index, black race, and previous preterm delivery.

Table 4
Validity testing and odds ratios for increasing Bishop Score and decreasing Cervical Score at 22–24 weeks for SPTD <35 weeks' gestation

	Bishop Score			Cervical Score		
	≥4*	≥5	≥6	≥7	<1.5*	<1.0
N (%) positive for test	289 (9.9)	83 (2.9)	26 (0.9)	8 (0.3)	82 (2.8)	21 (0.7)
Sensitivity%	27.6	11.8	7.9	3.9	13.4	6.3
95 CI	20.0–36.2	6.8–18.7	3.8–14.0	1.3–8.9	8.0–20.6	2.8–12.0
Specificity %	90.9	97.6	99.4	99.7	97.7	99.5
[False positive %]	[9.1]	[2.4]	[0.6]	[0.1]	[2.3]	[0.5]
Specificity 95% CI	(89.8–91.9)	96.9–98.1	99.1–99.7	99.7–100	97.0–98.2	99.2–99.8
Positive predictive value %	12.1	18.1	38.5	62.5	20.7	38.1
Adjusted + Odds Ratio	2.9	3.9	14.6	31.9	4.8	8.0
95% CI	1.9–4.5	2.1–7.3	5.9–36.2	6.6–154.4	2.6–8.7	2.9–21.9
Area under ROC curve	0.66			0.61		
95% CI	(0.61–0.71)			(0.56–0.67)		

* Selected diagnostic thresholds for SPTB <35 weeks at the 22–24 week examination

CI - confidence interval

+ Adjusted for body mass index, black race, and previous preterm delivery.

‡ Unable to compute due to a cell with zero patients.

Table 5

Validity testing and odds ratios for an increasing Bishop Score and a decreasing Cervical Score at 26–29 weeks for SPTD <35 weeks' gestation

	Bishop Score			Cervical Score		
	≥5*	≥6	≥7	<1.5*	<1.0	<0
N (%) positive for test	199 (7.8)	71 (2.8)	19 (0.8)	147 (5.8)	59 (2.3)	14 (0.6)
Sensitivity %	32.1	17.9	7.1	35.7	23.8	6.0
95% CI	22.4–43.2	10.4–27.7	2.7–14.9	25.6–46.9	15.2–34.3	2.0–13.3
Specificity %	93.0	97.7	99.5	95.2	98.4	99.6
[False positive %]	[7.0]	[2.3]	[0.5]	[4.8]	[1.6]	[0.4]
Specificity 95% CI	91.9–94.0	97.0–98.2	99.1–99.7	94.3–96.0	97.8–98.9	99.3–99.8
Positive predictive value %	13.6	21.1	31.6	20.4	33.9	35.7
Adjusted+ Odds Ratio	4.4	5.7	6.6	8.2	13.5	5.6
95% CI	2.7–7.4	2.9–11.1	2.2–19.4	4.9–13.6	7.1–25.5	1.7–18.5
Area under ROC curve		0.68			0.68	
95% CI		0.62–0.75			0.61–0.75	

* Selected diagnostic thresholds for SPTD <35 weeks at the 26–29 week examination

CI - confidence interval

+ Adjusted for body mass index, black race, and previous preterm delivery.

Table 6

Comparison of sonographic and digital cervical assessment at 26–29 weeks for prediction of SPTD <35 weeks' gestation

	Sonographic Assessment		Digital Assessment
	Cervical Length ≤20mm	Funneling Present	Cervical Score <1.5
Sensitivity (%)	32	32	36
Specificity (%)	95	91	95
Positive predictive value (%)	17	11	20
Negative predictive value (%)	98	98	98