Transvaginal ultrasonography compared with Bishop score for predicting cesarean section after induction of labor

Parvin Bastani
Kobra Hamdi
Fatemeh Abasalizadeh
Parisa Pourmousa
Fatemeh Ghatrehsamani
Women’s Reproductive Health Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Background: Reproductive health researchers are interested in finding better methods for predicting an unwanted type of delivery after induction of labor. The aim of this study was to compare the value of transvaginal ultrasonography findings and the Bishop score in predicting cesarean section after induction of labor.

Methods: Two hundred women with singleton pregnancies undergoing induction of labor at 37–42 weeks were enrolled in this prospective study. Transvaginal investigation was done for all participants prior to induction. To compare the predictive value of the methods, receiver-operating characteristic (ROC) curves were plotted and equality of the area under curve (AUC) was tested.

Results: The mean age of the participants was 29.9 years, mean gestational age was 39.6 weeks, and mean gravid was 1.5. The AUC calculated for Bishop score was 0.39 (95% confidence interval [CI] 0.3–0.48). The AUC for cervical length measured by ultrasonography was 0.69 (95% CI 0.6–0.77). The AUC for the posterior cervical angle measured by ultrasonography was 0.38 (95% CI 0.29–0.47). Testing equality of the ROC curves for these three methods showed the ROC for cervical length to be statistically different from both Bishop score and posterior cervical angle (P<0.001). However, the difference in ROC area compared between Bishop score and posterior cervical angle was not statistically significant.

Conclusion: Based on our findings and available information in the literature, it seems that cervical length measured by transvaginal ultrasonography has the potential to replace the traditional Bishop score, provided that such a facility is available when needed.

Keywords: induction of labor, Bishop score, transvaginal ultrasonography, cesarean section, diagnostic value

Introduction
Approximately 20% of pregnant women undergo induction of labor around term. However, not all of these inductions result in vaginal delivery, and some result in emergency cesarean sections. Compared with spontaneous onset of labor, induction of labor is complicated by a higher rate of cesarean section, occurring in one-fifth of women who undergo induction. An issue of interest for reproductive health researchers is finding better methods for predicting an unwanted type of delivery after induction of labor. The Bishop score is an index known to predict outcome of labor, but it does not provide satisfactory results, due to a low predictive value, especially in predicting cesarean section. Transvaginal ultrasonography is a known objective method for assessing cervical length. It is not too sophisticated a procedure to be done in obstetric units and clinics. Attempts have been made in the past to use transvaginal ultrasonography
for prediction of type of delivery, and efforts have been made to explore if its predictive value is higher than that of the Bishop score or not. However, controversial results have been published, and not enough evidence is available to consider it as a strong alternative to the Bishop score. The aim of this study was to compare the predictive value of transvaginal ultrasonographic findings and the Bishop score for cesarean section as an outcome of induction of labor.

Methods
Two hundred pregnant women admitted for induction of delivery were enrolled in this prospective study. They were hospitalized at Alzahra University Hospital in Tabriz, northwest Iran, in 2007–2009. Women with pregnancies at 37–42 weeks’ gestational age who agreed to participate were considered eligible for the study. The study protocol was approved by the ethics committee at Tabriz University of Medical Sciences. Inclusion criteria were a singleton pregnancy, cephalic presentation, gestational age more than 37 weeks, and a live pregnancy. To prevent iatrogenic premature rupture of membranes, those with cervical dilatation $3$ cm were excluded from the study. Other exclusion criteria were labor pains having already started and anencephaly.

Transvaginal investigation was done for all participants prior to induction of delivery. Cervical length and posterior cervical angle were measured by transvaginal sonography, using an Aloka-350 ultrasound machine and a $15$ mHz convex probe. Posterior cervical angle, defined as the angle between the cervical canal and the posterior uterine wall, was measured using the trace function of the ultrasound machine.$^3$ Bishop score was assessed by digital examination prior to induction.

Statistical analysis
Data were entered into the computer and analyzed using the STATA 11 statistical software package (STATA Corporation, College Station, TX). Descriptive sensitivity and specificity statistics were calculated. To compare the predictive value of the methods, receiver-operating characteristic (ROC) curves were plotted and equality of the area under the curve (AUC) was tested. Correct classification rates were calculated by STATA, based on frequency of discordant pairs.

Results
Mean age ($\pm$ standard deviation [SD]) of the participants was $29.9$ ($\pm 5.6$) years, mean gestational age was $39.6$ ($\pm 1.4$) years, and mean gravid was $1.5$ ($\pm 0.8$). Up to $10\%$ of the participants declared a previous abortion history, up to $10\%$ were illiterate, and two-thirds were nulliparous.

Mean Bishop score was $4.1$ ($\pm 1.8$). Mean cervical length was $18.1$ ($\pm 5.5$) cm. Mean posterior cervical angle was $116.5$ ($\pm 12$) degrees. Of $200$ deliveries, $143$ ($71.5\%$) were vaginal and $57$ ($28.5\%$) were by cesarean section. Fifty-seven percent of the neonates were male. The first-minute Apgar score in half of the neonates was $\geq 9$, $6$ in one, and $8$ in $18$ cases. Except for two babies with a $5$-minute Apgar score $= 8$, the babies had higher scores at the fifth minute.

Table 1 presents sensitivity, specificity, and correct classification rates for the different methods of predicting cesarean section. The wide range of cut points was truncated to those producing more acceptable sensitivity and specificity.

The ROC curves for Bishop score, cervical length, and posterior cervical angle for predicting cesarean section versus vaginal delivery are compared in Figures 1 and 2. The AUC for Bishop score was $0.39$ (95% CI $0.3$–$0.48$). The AUC for cervical length measured by ultrasonography was $0.69$ (95% CI $0.6$–$0.77$). The AUC for posterior cervical angle measured by ultrasonography was $0.38$ (95% CI $0.29$–$0.47$). The ROC for cervical length was significantly different from both Bishop score and posterior cervical angle ($P < 0.001$). However, the difference for AUC compared for Bishop

<table>
<thead>
<tr>
<th>Cut point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Correctly classified (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bishop score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\geq 2$</td>
<td>80.7</td>
<td>4.2</td>
<td>26</td>
</tr>
<tr>
<td>$\geq 3$</td>
<td>71.9</td>
<td>11.2</td>
<td>28.5</td>
</tr>
<tr>
<td>$\geq 4^*$</td>
<td>57.9</td>
<td>28.7</td>
<td>37</td>
</tr>
<tr>
<td>$\geq 5$</td>
<td>36.8</td>
<td>47.6</td>
<td>44.5</td>
</tr>
<tr>
<td>$\geq 6$</td>
<td>12.3</td>
<td>85.3</td>
<td>64.5</td>
</tr>
<tr>
<td>Cervical length</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\geq 16$</td>
<td>78.9</td>
<td>39.9</td>
<td>51</td>
</tr>
<tr>
<td>$\geq 17$</td>
<td>75.4</td>
<td>53.2</td>
<td>59.5</td>
</tr>
<tr>
<td>$\geq 18$</td>
<td>68.4</td>
<td>58</td>
<td>61</td>
</tr>
<tr>
<td>$\geq 19^*$</td>
<td>66.7</td>
<td>65</td>
<td>65.5</td>
</tr>
<tr>
<td>$\geq 20$</td>
<td>61.4</td>
<td>69.9</td>
<td>67.5</td>
</tr>
<tr>
<td>$\geq 21$</td>
<td>50.9</td>
<td>80.4</td>
<td>72</td>
</tr>
<tr>
<td>$\geq 22$</td>
<td>50.9</td>
<td>82.5</td>
<td>73.5</td>
</tr>
<tr>
<td>Posterior cervical angle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\geq 114$</td>
<td>56.1</td>
<td>21.7</td>
<td>31.5</td>
</tr>
<tr>
<td>$\geq 115$</td>
<td>52.6</td>
<td>22.4</td>
<td>31</td>
</tr>
<tr>
<td>$\geq 116$</td>
<td>49.1</td>
<td>30.1</td>
<td>35.5</td>
</tr>
<tr>
<td>$\geq 117^*$</td>
<td>45.6</td>
<td>33.6</td>
<td>37</td>
</tr>
<tr>
<td>$\geq 118$</td>
<td>42.1</td>
<td>37.1</td>
<td>38.5</td>
</tr>
<tr>
<td>$\geq 119$</td>
<td>35.1</td>
<td>42.7</td>
<td>40.5</td>
</tr>
<tr>
<td>$\geq 120$</td>
<td>35.1</td>
<td>53.8</td>
<td>48.5</td>
</tr>
</tbody>
</table>

Note: *Cut points giving the most appropriate test values.
Transvaginal US versus Bishop score

score and posterior cervical angle was not statistically significant.

Bishop score was correlated with posterior cervical angle ($r = 0.24, P < 0.001$). It was inversely correlated with cervical length ($r = -0.51, P < 0.001$) and interval until induction of delivery ($r = -0.21, P < 0.01$). The comparison scatter plots are given in Figure 3. Induction interval was not found to be correlated with posterior cervical angle, but was correlated with cervical length ($r = -0.22, P < 0.01$).

Discussion

Imaging versus clinical assessment has always been a challenge for choosing diagnostic or prognostic measures, sometimes showing simple clinical prediction scores to be as good as sophisticated imaging facilities.3,4 In the present study, cervical length measured by transvaginal ultrasonography was found to be a better predictor of an unsuccessful labor leading to cesarean section than the Bishop score.

However, the ability of the posterior cervical angle to predict cesarean section was not much different from that of the Bishop score.

Several previous studies have compared ultrasonographic findings with the Bishop score in assessing the success of induction. Most of these studies have used smaller sample sizes than ours, and few have had larger samples. Regardless of slight differences in methodology, all the larger-scale studies, like ours, have found cervical length to be a better predictor than Bishop score.1,2,5–13 Few have found cervical length to be the best predictor, few have prioritized the Bishop score,14–16 and two studies have not observed a difference in their predictive value.17,18 Studies suggesting Bishop score as the best predictor have not compared the AUC or provided sensitivity and specificity measures for given cut points. None of the other studies mentioned have reported AUC for cervical length to be $0.66$ and $0.89$, and the AUCs reported for Bishop score vary widely, sometimes being as low as 0.46.

We did not find cervical angle to be a good predictor of delivery type. Novakov-Mikic et al reported findings somewhat similar to ours.19 However, they investigated the anterior cervical angle, the predictive value of which was even lower than ours. Eggebo et al found the posterior cervical angle to be a predictor of success of induction, but in a regression model rather than an assessment of predictive value.19

Conclusion

Based on our findings and the information available in the literature, it seems that cervical length measured by transvaginal ultrasonography has the potential to replace the traditional Bishop score when such a facility can be made easily available. However, prior to general
recommendation of this approach, cost-effectiveness studies will be necessary.

Disclosure
The authors report no conflicts of interest in this work.

References