An ultrasound prediction model for probability of vaginal delivery in induction of labor

Poonam Garg1, Maria Dolores Gomez Roig1, Aprajita Singla1

1Department of Obstetrics and Gynecology, Chaitanya Hospital, Chandigarh, India
2Department of Obstetrics and Gynecology, Sant Joan De Deu Hospital, Barcelona, Spain
3Department of Public Health, Panjab University, Chandigarh, India

Abstract

Objective: Our aim was (i) to evaluate a pre-induction ultrasound score for prediction of vaginal delivery and compare it with the Bishop score in term nulliparous women, and (ii) to formulate a prediction model to calculate probability of vaginal delivery for clinical use.

Methods: Ninety-six nulliparous women between 36–41 weeks gestation were recruited. All subjects fulfilled the inclusion criteria of a live singleton pregnancy, vertex presentation, intact amniotic membrane, in the absence of active labor with no contraindication to vaginal delivery. The patients were assessed by our ultrasound score comprising of 3 cervical and 2 fetal head parameters. These parameters were fetal head position, fetal head - symphysis pubis distance relation, cervical length, funneling and posterior cervical angle. Each parameter was scored from 0–2, with a maximum score of 10. A second obstetrician blinded to the sonographic findings assessed the modified Bishop score. SPSS 20 was used for ROC curves plots and calculation of area under curve. Binary logistic regression model was prepared and probability of vaginal delivery for various scores was calculated.

Results: Out of 91, 61 (67%) achieved active phase of labor and 54 (59%) had vaginal delivery. Our pelvic ultrasound score showed better sensitivity and specificity in comparison to the Bishop score. At a cut-off of ≥5, the ultrasound score showed sensitivity of 79.3%, specificity of 75.8%, whereas, the Bishop score showed sensitivity of 66.7% and specificity of 44.2%. Binary logistic regression model predicted 78.0% of the events correctly.

Conclusion: Our study shows that “Garg ultrasound score” can predict success of induction of labor in nulliparous women. This proposed pelvic ultrasound score, if validated in larger multicenter studies, could help clinicians provide evidence-based counselling for predicting probability of vaginal delivery. This in turn, may allow women make a more informed decision before undergoing induction of labor.

Keywords: Induction of labor, Bishop score, Garg ultrasound score, induction success, prediction model.

Original Article
Perinatal Journal 2019;27(3):161–168
©2019 Perinatal Medicine Foundation
Introduction
Labor is a complex physiological process, and we have limited understanding of factors that initiate the process of labor. Induction of labor is an obstetric intervention undertaken when continuation of pregnancy is thought to be associated with maternal or fetal risk, with an aim to have a vaginal delivery. While it is one of the most common obstetrical procedures,1,2 our ability to predict success of induction is limited.

Induction of labor has been found to have a major impact on the labor experience of women. In this era of positive labor and shared decision making, this needs serious introspection. It shows that clinicians need better tools to predict success in induction of labor, so that they can customize their counseling based on individual patient characteristics.3,4 Since 1941, many scoring systems have been suggested and till date Bishop score is universally the most accepted one. However, this is a subjective method with high inter and intra observer variability, uncomfortable and less precise.5-8 So obstetricians need an objective, painless, easy and precise method.

In the last 30 years, lot of work has been done to evaluate the role of ultrasound in prediction of success of induction. Sonographic cervical length was the first ultrasound parameter to be studied; however, the results are conflicting with no consensus on appropriate cut-off value.9 In recent years’ various other ultrasound parameters like posterior cervical angle, funneling, wedging pattern, fetal head position, fetal head - perineum distance, fetal head - symphysis pubis distance have been evaluated either individually or in combination and compared with Bishop score.9-11 Many authors formulated scoring systems based on combination of ultrasound and clinical parameters and compared with Bishop score.12-14 In spite of large volume of work and acceptability of superiority of ultrasound over digital examination still we have not been able to put it into clinical practice. This has prompted us to design and study Garg scoring system based on five ultrasound parameters – 3 cervical and 2 fetal head. This is in consonance with Bishop score which has 4 cervical and 1 fetal head parameter. We have selected these parameters based on results of previous studies and have been found to be easily measurable and have shown favorable correlation in prediction of successful induction: cervical length, posterior cervical angle, funneling, fetal head position, fetal head symphysis pubis relation.

The present study is to compare the five parameters of modified Bishop score15,16 and five parameters of Garg ultrasound score to predict successful labor induction. Our aim was (i) to evaluate a pre-induction ultrasound score for the prediction of vaginal delivery and compare it with the Bishop score in term nulliparous women, and (ii) to formulate a prediction model to calculate probability of vaginal delivery for clinical use.

Methods
This was a quasi-experimental study which included 96 women admitted for induction of labor at Chaitanya hospital, Chandigarh, India between December 2017 to Jan 2019. Inclusion criteria was singleton nullipara between 36–41 weeks with live fetus in vertex presentation with intact amniotic membranes and no signs of labor. Exclusion criteria included multiple pregnancy, malpresentation, previous scarred uterus, abnormal placentation, fetus with congenital anomalies, fetal intrauterine death, cervical cerclage in the present pregnancy. All women gave consent to participate in the study. This study received approval of ethics committee of hospital.

On admission a detailed history was followed by general and systemic examination. The gestational age was reconfirmed based on the date of last menstrual period and ultrasound measurement at or before 12 weeks of gestation. All ultrasounds were performed by one investigator using Voluson S6 Ultrasound machine (GE Medical Systems, Milwaukee, WI, USA). Immediately before induction and bladder emptying determination of the fetal occipital position and fetal head - symphysis pubis relation (FHPR) was done by transabdominal ultrasound. A transvaginal sonography was done to assess the cervical funneling, cervical length and posterior cervical angle. A printout of all ultrasounds was taken and measurement was done using scale and protractor and recorded in performa. Subsequently a second obstetrician blinded to the sonographic findings assessed the Bishop score and findings were recorded in a separate performa.

Sonographic measurements of cervical length, posterior cervical angle, fetal head position were measured
according to Rane et al.\[10\] funneling was according to Chung et al.\[11\] For FHPRT determination, the convex probe was kept vertically on pubic region so as to visualize the symphysis pubis and fetal head simultaneously. A line perpendicular and starting from superior margin of symphysis pubis was drawn toward fetal head and checked if the distance between this line and fetal head was measurable, not measurable or touching (as shown in Figs. 1a–c). Each parameter was scored from 0–2, with maximum score of 10 (Table 1). Modified Bishop score was assessed and scored as per Table 2.

For induction of labor, cervical ripening was done with buccal Tab misoprostol 25 mcg, maximum of three doses at 4-hour interval until uterine contractions reached a frequency of three in 10 mins. Oxytocin if required was started in escalating doses. Continuous electronic fetal heart rate monitoring and tocodynamometry was used in all women. Success of induction of labor was defined as the vaginal delivery. Failed induction was defined as inability to achieve active phase of labor corresponding to cervical dilatation of $\geq 5$ cm within 8 hours of initiating oxytocin. Non-progress of labor was defined as no cervical dilatation after active phase of labor for at least 2 hours and/or no decent of fetal head during second stage of labor for at least 2 hours despite of adequate uterine contraction. Failed induction, non-progress of labor and fetal distress were considered as an indication for cesarean delivery.

Descriptive analysis was done in form of numbers, percentages, mean and standard deviation represented in form of tables and figures. Student’s t-test and Pearson’s chi-squared test were applied to check significance of results. Sensitivity, specificity, false positive rate and correct classification rate (accuracy) were calculated for vaginal delivery. ROC (receiver operating characteristic) curves were plotted for most suitable cut-off points 4, 5 and 6. The data collected were analyzed using MS Excel (Microsoft Corp., Redmond, WA, USA) 2007 and SPSS 20 (SPSS Inc., Chicago, IL, USA).

Results

A total of 96 women were enrolled. However, 5 women did not complete the trial. Therefore, 91 were included and evaluated (Fig. 2). Indications for labor induction were: cholestasis of pregnancy (n=24), gestational diabetes (n=15), gestational hypertension (n=4), intrauterine growth retardation (n=7), post datism (n=23) and elective (n=18). Of 91 women, 61 (67%) went into active phase of labor. 54 (88.5%) delivered vaginally and 7
(11.5%) required cesarean – 5 for non-progress of labor and 2 for fetal distress. Of 30 (33%) women who did not achieve active phase of labor and underwent cesarean, 15 had failed induction and 15 developed fetal distress. Regarding mode of delivery, 54 (59.4%) out of 91 enrolled women achieved vaginal delivery and 37 (40.6%) required cesarean – 17 for fetal distress, 15 for failed induction and 5 for non-progress of labor.

The characteristics of 91 patients are given in Table 3. The mean age of enrolled women was 29.87±3.29 (range: 21–40) years. Mean body mass index (BMI) was 29.09±4.43 (range: 19.7–46.1). Mean gestation was 269.2±7.96 (range: 256–282) days. Table 4 gives comparison of sensitivity, specificity and other diagnostic parameters at different cut-offs of 4, 5 and 6. Fig. 3 gives comparison of ROC curves for prediction of mode of delivery.

The cut-off value of 5 for Garg ultrasound score [sensitivity 79.3%, specificity 75.8%] and cut-off value of 4 for Bishop score [sensitivity 69.0%, specificity 55.6%] produce best combination. ROC was better for Garg ultrasound score than Bishop score with area under the curve (AUC) .850 in comparison to AUC of .622 (Table 5). Garg ultrasound score shows highly sig-
significant p-value. A model of binary logistic regression, with equation formula: probability = 1/1 + e^{-x}, was designed that included total Garg ultrasound score as variable.

The following equation was obtained: \( p=\frac{1}{1 + \text{EXP}(-5.739-1.294\times\text{USGS})} \) where \( p \) is the probability of vaginal delivery and USGS is total Garg ultrasound score obtained. Values equal or greater than 0.5 predicted vaginal delivery. The model predicted 78.0% of the events correctly (Table 6).

Table 7 shows calculated probabilities for each calculated value of Garg ultrasound score so that obstetrician can use it easily in practice. Neonatal outcomes are shown in Table 8. Mean baby weight at the time of delivery was 2957 (range: 1890–3870) g. Mean 1-minute Apgar score was 7.8 (range: 4–9) and at 5-minute Apgar score was 8.9 (range: 6–9). Out of 91 babies, 82 (91%) were given to their mothers and only 8 (9%) required specialized neonatal intensive care. None of our enrolled women developed complications associated with induction of labor like uterine hyper stimulation and rupture of uterus. There was no maternal and neonatal mortality.

**Discussion**

This study shows that Garg ultrasound scoring system has sensitivity similar to Bishop score in the prediction of vaginal delivery (79.3% vs 69%). However, Garg ultra-

<table>
<thead>
<tr>
<th>Scoring method</th>
<th>Cut-off scores</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>False positive rate (1-specificity) (%)</th>
<th>Correct classification rate (Accuracy) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bishop score</td>
<td>≥4</td>
<td>69.0</td>
<td>55.6</td>
<td>44.4</td>
<td>63.7</td>
</tr>
<tr>
<td></td>
<td>≥5</td>
<td>66.7</td>
<td>44.2</td>
<td>55.7</td>
<td>51.6</td>
</tr>
<tr>
<td></td>
<td>≥6</td>
<td>60.0</td>
<td>40.7</td>
<td>59.3</td>
<td>41.7</td>
</tr>
<tr>
<td>Ultrasound score</td>
<td>≥4</td>
<td>71.2</td>
<td>88.9</td>
<td>11.1</td>
<td>74.7</td>
</tr>
<tr>
<td></td>
<td>≥5</td>
<td>79.3</td>
<td>75.8</td>
<td>24.2</td>
<td>78.0</td>
</tr>
<tr>
<td></td>
<td>≥6</td>
<td>93.1</td>
<td>56.5</td>
<td>43.5</td>
<td>68.1</td>
</tr>
</tbody>
</table>

Table 5. Area under curve in ROC analysis according to mode of delivery.

<table>
<thead>
<tr>
<th>Test result variable(s) according to mode of delivery</th>
<th>AUC (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garg ultrasound score (range)</td>
<td>.850 (.772–.927)</td>
<td>.000</td>
</tr>
<tr>
<td>Bishop score (range)</td>
<td>.622 (.502–.743)</td>
<td>.048</td>
</tr>
</tbody>
</table>

AUC: area under the curve; CI: confidence interval.

Table 6. Predictive model.*

<table>
<thead>
<tr>
<th>Observed mode of delivery</th>
<th>Expected mode of delivery</th>
<th>Success (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSCS</td>
<td>Mode of delivery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LSCS</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>NVD</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Overall percentage</td>
<td></td>
</tr>
</tbody>
</table>

LSCS: lower segment cesarian section; NVD: normal vaginal delivery. *The cut-off value is .500.
sound scoring system has a clearly significant higher specificity (75.8% vs 55.6%) for the vaginal delivery and hence is superior to Bishop score. Though our ultrasound scoring system has shown statistical significance at scores of 4, 5 and 6, we propose cut-off of 5 since it has shown a combination of maximum sensitivity and specificity and accuracy.

The success of induction depends on favorability of cervix and fetal head position which till date is assessed by manual examination and scored as Bishop score. However, clinically, it has shown to be inherently subjective with high inter and intra observer variability and variable sensitivity and specificity. To overcome this limitation, a number of researchers have studied accuracy of one or more ultrasound parameters either alone or in combination with Bishop score and/or maternal characteristics for prediction of successful induction of labor. However, none of these have so far gained widespread acceptability and clinical application because of conflicting results. The five parameters of ultrasound scoring system proposed by us have individually shown to have good accuracy in various studies. Out of these 5 parameters, 3 assess cervical favorability and 2 assess fetal head position and location in relation to symphysis pubis.

Table 7. Calculated probabilities for each Garg ultrasound score.

<table>
<thead>
<tr>
<th>Total ultrasound score</th>
<th>Estimated probability of vaginal delivery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>2</td>
<td>4.1</td>
</tr>
<tr>
<td>3</td>
<td>13.5</td>
</tr>
<tr>
<td>4</td>
<td>36.3</td>
</tr>
<tr>
<td>5</td>
<td>67.5</td>
</tr>
<tr>
<td>6</td>
<td>88.3</td>
</tr>
<tr>
<td>7</td>
<td>91.8</td>
</tr>
<tr>
<td>8</td>
<td>96.6</td>
</tr>
<tr>
<td>9</td>
<td>98.6</td>
</tr>
</tbody>
</table>

Ultrasound measurement of cervical length has been the most studied single parameter which has been compared with Bishop score for prediction of successful induction of labor. Due to conflicting results in its prediction of success and wide variation in cut-off points when used alone is not useful in prediction of vaginal delivery. Our scoring system has five parameters including cervical length and has shown statistically significant results. A number of authors have studied and a few have proposed scoring systems consisting of a combination of ultrasound parameters, one or more Bishop score parameters and/or maternal characteristics. Rane et al. and Keepanasseril et al. have shown good sensitivity and specificity but their results cannot be compared to our study because their patient population included multiparous women. We feel that maternal characteristics should not be included in scoring system since they remain unchanged regardless of whether the patients are induced or come in spontaneous labor. Similarly, the scoring system proposed by Eggebo et al. includes digitally measured dilatation of cervix which again brings the limitation of subjectivity and patient discomfort.

Bajpai et al. formulated ultrasound scoring system with parameters which matched the components of Burnet modified Bishop score including cervical length, funnel length, funnel width, position of cervix and distance of presenting part to external os. At cut-off of 4, they achieved sensitivity of 64.52% and specificity of 85.71% for Bishop score and sensitivity of 77.42% and specificity of 92.86% of ultrasound score respectively for prediction of entering into active phase of labor. However, these scoring system parameters, especially funnel length and width, have difference of only 5 mm which adds subjectivity to the measurement. Their study population includes both nullipara and multipara with a greater number of multiparous women.

The strengths of the Garg ultrasound score include: (1) it has five parameters which take into account cervical as well as fetal head status thus providing a more complete assessment similar to Bishop score. (2) We have well defined protocol for measurement of each parameter. All these parameters are easily measurable, do not require high degree of skill, can be performed quickly and provide an objective assessment. (3) Ultrasound is well tolerated and hence more accept-
able to patients as well as clinicians. (4) Our prediction score can be handy for obstetrician in clinical practice.

The potential limitations of the present study are as follows: (1) It has a small sample size so validation through larger multicenter study is required. (2) It is studied in only nulliparous women before induction of labor, so we need to test diagnostic accuracy in other groups like multiparous women, spontaneous labor, VBAC and preterm labor. (3) Cervix undergoes changes with the increasing weeks of gestation, and our study included patients from 36–41 weeks, so it is better to convert cervical length into MOMS. (4) The vaginal delivery depends on fetal head circumference and maternal pelvic parameters like inter ischial diameter and subpubic angle.[21] The inclusion of these parameters may improve the accuracy. (5) The confounding variable effect such as epidural analgesia needs to be evaluated. (6) Our cesarean rate is high (40%) in comparison to the expected cesarean rate of 25–30% according to Robson classification. (7) The use of misoprostol as induction agent can affect the results due to its association with fetal distress. So, we need to validate Garg ultrasound score in women using other induction agents (foleys, dinoprostone).

Our study has shown high sensitivity and specificity, and we suggest that Garg ultrasound scoring system has an excellent scope of being used in clinical practice for prediction of success of induction of labor. It can help clinicians in providing evidence-based counseling to pregnant women for informed shared decision making.

Conclusion

Through this study, we propose Garg ultrasound scoring system which can predict success of induction of labor with greater specificity as compared to the traditional Bishop score. This scoring system is highly objective, reproducible, easy to perform and does not include digital vaginal examination which is subjective and painful for patients. With this scoring system, if validated in more diverse population in larger multicentric studies, clinicians should be able to provide individualized counseling and help women to make more informed decision.

Conflicts of Interest: No conflicts declared.

References