Transvaginal Ultrasonographic Cervical Measurement in Predicting Failed Labor Induction and Cesarean Delivery for Failure to Progress in Nulliparous Women

The aim of this study was to evaluate the value of transvaginal sonographic cervical measurement in predicting failed labor induction and cesarean delivery for failure to progress in nulliparous women. One hundred and sixty-one women scheduled for labor induction underwent transvaginal ultrasonography and digital cervical examinations. Logistic regression demonstrated that cervical length and gestational age at induction, but not the Bishop score, significantly and independently predicted failed labor induction. According to the receiver operating characteristic curves analysis, the best cut-off value of cervical length for predicting failed labor induction was 28 mm, with a sensitivity of 62% and a specificity of 60%. In terms of the likelihood of a cesarean delivery for failure to progress as the outcome variable, logistic regression indicated that maternal height and birth weight, but not cervical length or Bishop score, were significantly and independently associated with an increased risk of cesarean delivery for failure to progress. Transvaginal sonographic measurements of cervical length thus independently predicted failed labor induction in nulliparous women. However, the relatively poor predictive performance of this test undermines its clinical usefulness as a predictor of failed labor induction. Moreover, cervical length appears to have a poor predictive value for the likelihood of a cesarean delivery for failure to progress.

Key Words : *Cervical Length; Bishop Score; Labor Induction; Cesarean Delivery for Failure to Progress*

Kyo Hoon Park

Department of Obstetrics and Gynecology, Seoul National University College of Medicine, Seoul National University Bundang Hospital, Seongnam, Korea

Address for correspondence
Kyo Hoon Park, M.D.
Department of Obstetrics and Gynecology, Seoul National University Bundang Hospital, 300 Gumi-dong, Bundang-gu, Seongnam 463-707, Korea
Tel : +82.31-787-7252, Fax : +82.31-787-4054
E-mail : pkh0419@snubh.org

INTRODUCTION

Induction of labor is a common and essential element of contemporary obstetric practice with an incidence of approximately 20% of pregnancies (1). Although induction of labor is a safe and efficacious method of vaginal delivery, it still carries considerable risk of cesarean delivery compared with spontaneous labor, and this risk is significantly influenced by the status of the cervix at the time of labor induction (2-4). Therefore, several attempts have been made to assess cervical status using Bishop score and transvaginal ultrasound findings.

Transvaginal ultrasonography is a well known objective technique for assessing cervical length and the morphologic characteristics of the internal cervical os (5, 6). Recently this assessment, given the established role of transvaginal ultrasonography in the management of patients at risk of preterm delivery (7), has been proposed as a means of predicting the outcome of induction. However, the results of studies that have evaluated the accuracy of transvaginal ultrasound for the prediction of the outcome of induction remain contro-

versial. Some investigators have reported that transvaginal ultrasonographic assessments of the cervix are of value for predicting the outcome of induction (8-10), whereas others have argued against it (11-14). The reason for these discrepant results may relate to the outcome variables used for analysis. In most studies, investigators have used the likelihood of vaginal delivery or the time interval between induction and delivery as primary outcome variables. However, the use of these parameters as outcome variables obviously affects conclusions because they are influenced by factors other than cervical status, for example, by parity, birth weight, maternal stature, and indications for cesarean delivery (2, 15, 16).

The active phase of labor consistently begins at a cervical dilatation of 3 to 4 cm or more, and its onset is considered to be a reasonably reliable diagnostic threshold of active labor because it demarcates the transition from desultory progression in the latent phase to more rapid cervical dilatation. In addition, the American College of Obstetricians and Gynecologists has proposed that the performance of cesarean delivery for dystocia be specifically indicated only in women in the active phase of labor (17). Therefore, we focused on the

final outcomes of labor induction using two indices, a lack of progression into the active phase of labor and labor abnormalities during the active phase and the second stage of labor. The purpose of this study was to evaluate the value of transvaginal sonographic cervical measurement for predicting failed labor induction and cesarean delivery for failure to progress in nulliparous women.

MATERIALS AND METHODS

Study design

This prospective observational study enrolled 161 consecutive women admitted for induction of labor to Seoul National University Bundang Hospital (Seongnam, Korea) between March 2004 and July 2006. The inclusion criteria were: 1) nulliparous patients; 2) singleton pregnancy; 3) live fetus with vertex presentation; 4) intact amniotic membranes; 5) ≥ 37 weeks gestation; 6) absence of preeclampsia (this criterion was developed because prior studies had shown that preeclampsia or use of magnesium sulfate is independently associated with the failed labor induction and cesarean delivery (2, 18); 7) absence of active labor; 8) an initial cervical examination showing ≤ 2 cm dilatation and $\leq 50\%$ effacement; 9) no history of previous uterine surgery, and 10) no contraindication to vaginal delivery. The decision to induce labor was made by attending physicians. Labor induction was defined as stimulation of uterine contractions before the spontaneous onset of labor for the purpose of accomplishing delivery in women with obstetric, medical, or fetal complications of pregnancy. This prospective study was approved by the institutional review board of our institution (Seoul National University Bundang Hospital, Seongnam, Korea). Informed consent was obtained prior to enrollment in the study.

Transvaginal ultrasound and digital clinical assessment of the cervix

On admission to the labor and delivery floor, transvaginal ultrasonographic assessment of cervical length and presence of funneling was performed by the author using an Envisor (Philips Medical System, Netherlands) ultrasound machine with a 6.0 MHz transducer. The women were asked to empty their bladder and were placed in the dorsal lithotomy position. The probe was gently placed at the anterior fornix of the vagina to obtain a sagittal view of the complete cervix, including the internal os, external os, and endocervical canal. The probe was slowly withdrawn until the image blurred, and then the insertion pressure was increased only enough to restore a clear cervical image. This cervical image was magnified to at least 75% of the screen, and the electronic markers were placed at the farthest points between the internal os and external os, then the cervical length was measured

as a straight line. The shortest of three measurements obtained was taken as the cervical length. Funneling was defined as a V- or U- shaped indentation of the internal os, and the measurement was taken from the apex of the funnel to the plane of the internal os. The Bishop score was determined (19) by the resident physician responsible for the induction who was not aware of the sonographic findings.

Induction protocol

Induction of labor was carried out according to the standard labor induction protocol used at the Seoul National University Bundang Hospital. External cardiotocography was performed to assess fetal well-being and to confirm the absence of contraction. If the Bishop score was 4 or less, a 10-mg dinoprostone vaginal insert (Propess, Controlled Therapeutics Ltd, East Kilbride, U.K.) was placed transversely in the posterior fornix of the vagina at 9:00 p.m. to initiate cervical ripening before labor induction. After 9 hr, i.e., 6:00 a.m. on the following day, the dinoprostone vaginal insert was removed and an intravenous oxytocin infusion was started if regular uterine contractions and cervical change had not developed. Oxytocin was administered intravenously as a dilute solution using a constant infusion pump. The initial dose was set at 5.3 mU/min, and this was increased by one-half of the previous infusion rate every 30 min up to a maximum dose of 40 mU/min until regular painful uterine contractions ensued or labor progressed. Oxytocin infusion was continued for at least 12 hr. In patients who had not gone into active labor during this 12-hr period, labor was induced by repeating the same protocol (i.e., serial induction was performed). Fetal heart rates were monitored using a continuous electronic system in all patients. Amniotomy was not performed until the cervix became dilated by at least 3 cm and the vertex was engaged. One hundred and thirty-six patients received both prostaglandin and oxytocin; 14 were given oxytocin only and eleven patients were delivered after prostaglandin administration only.

Definition of failed induction and failure to progress

Failed induction was defined as an inability to achieve the active phase of labor corresponding to a cervical dilatation of ≥ 4 cm within 12 hr of initiating oxytocin (i.e., within 21 hr of the insertion of a dinoprostone vaginal insert). Failure to progress was defined as no cervical dilatation during the active phase of labor for at least 2 hr or no descent of the fetus's head during the second stage of labor for at least 1 hr despite adequate uterine contraction. This was considered an indication for cesarean delivery for failure to progress.

Statistical analysis

Univariate analysis was conducted with the Student t-test,

Mann-Whitney U test, or χ^2 test. Receiver operating characteristic (ROC) curves were constructed to describe the relationship between the sensitivity (true-positive rate) and the false-positive rate for the cervical length in the prediction of failed labor induction. Multiple logistic regression analysis was used to identify, among various clinical and ultrasonographic variables, which variables were independently associated with failed labor induction and cesarean delivery for failure to progress. Variables found to have a significant correlation or a tendency towards association with failed labor induction in univariate analysis ($p<0.25$) were entered into a logistic regression model. p values of <0.05 were considered statistically significant.

RESULTS

During the study period, 174 consecutive nulliparous women underwent labor induction. Of these, 13 preeclamptic women were excluded for not meeting the entry criteria, and a total of 161 women were included in the study. Indications for labor induction were: oligohydramnios (71 women); prolonged pregnancy of 41 completed weeks or greater (n=43); suspected fetal growth restriction (n=19); large for gestational age (n=9); gestational diabetes (n=7); non-reassuring non-stress test (n=4); chronic hypertension (n=2), and patient request for social reasons (n=6).

Vaginal delivery occurred in 127 (79%) women, and in 95 (59%) of these delivery was within 24 hr of induction. Cesarean delivery was performed in 34 (21%) women because of failure to progress (n=24), failure of induction (n=5), suspected fetal distress (n=3), and maternal request for social reasons or discomfort (n=2), and these 34 included 11 women in whom cesarean section was performed within 24 hr of induction for failure to progress (n=11). Data from 10 women who delivered by cesarean section for induction failure, sus-

pected fetal distress, or due to some social reasons during the latent phase of labor were censored and were not included in the analysis of the relationship between cesarean delivery for failure to progress and covariate.

Table 1 describes the clinical characteristics and obstetric outcomes of patients according to the success or failure of labor induction. Induction of labor was successful in 66% (106/161), and 55 women failed to enter the active phase of labor within 24 hr of induction. No significant differences were observed in the mean maternal age and weight, the prevalence of prostaglandin use, the prevalence of presence of funneling, or the mean birth weight between these two patient groups. However, the women who failed to induce labor had a lower median Bishop score at admission, a longer mean cervical length, and labor was induced at earlier gestational age than in those who induced labor successfully. Moreover, women who failed to induce labor had a significantly higher rate of cesarean delivery than did those who successfully induced labor.

Fig. 1 displays the ROC curves for sonographically measured cervical length for predicting failed induction. The curve constructed for sonographically measured cervical length was above the 45° line, indicating a significant relationship between this variable and failed induction (cervical length: area under the curve 0.678; SE 0.044; $p<0.0001$). The best cut-off value for predicting failed induction was 28 mm with a sensitivity of 62% and a specificity of 60%.

Multiple logistic regression analyses were performed to examine the relationship between failed labor induction and various clinical and ultrasonographic variables (Table 2). The variables analyzed included gestational age at induction, maternal age, Bishop score, cervical length, presence of fun-

Table 1. The clinical characteristics of study population

Characteristics	Induction of labor		p values
	Success (n=106)	Failure (n=55)	
Maternal age (yr)	30.6±3.6	29.9±3.5	NS
Maternal weight (kg)	68.1±9.6	68.1±8.8	NS
Gestational age at induction (weeks)	39.8±1.4	38.8±1.5	<0.001
Use of prostaglandin	91% (96/106)	93% (51/55)	NS
Birth weight (kg)	3.2±0.5	3.1±0.5	NS
Bishop score before induction	3 (0-7)	2 (0-7)	<0.01
Cervical length by ultrasound (mm)	25.5±6.4	30.2±7.3	<0.001
Presence of funneling	14% (15/106)	7% (4/55)	NS
Cesarean delivery	10% (11/106)	42% (23/55)	<0.001

Values are given as mean±SD, median (range) or % (n).
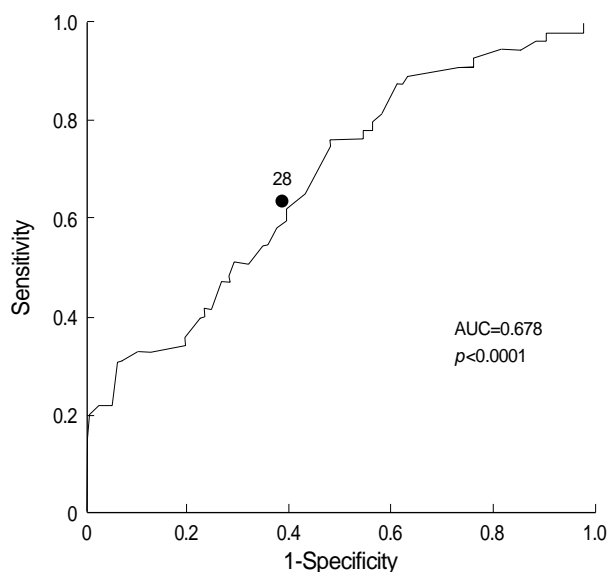NS, not significant.

Fig. 1. Receiver operating characteristic curves analysis for sonographically measured cervical length (mm) in predicting failed labor induction.

Table 2. The relationship between failed labor induction and independent variables analyzed by multiple logistic regression

	Odds ratio	95% Confidence interval	p values
Gestational age at induction	0.53	0.39-0.73	<0.001
Maternal age	0.96	0.86-1.08	NS
Cervical length by ultrasound	2.80	1.46-5.38	<0.005
Presence of funneling	0.83	0.21-3.25	NS
Bishop score at induction	0.85	0.67-1.07	NS
Birth weight	1.70	0.68-4.24	NS

NS, not significant.

Table 3. The relationship between cesarean delivery for failure to progress and independent variables analyzed by multiple logistic regression

	Odds ratio	95% Confidence interval	p values
Gestational age at induction	0.67	0.35-1.28	NS
Maternal weight	1.07	0.99-1.16	NS
Maternal height	0.77	0.65-0.90	<0.005
Epidural analgesia	3.12	0.71-13.71	NS
Maternal age	1.23	0.99-1.52	0.06
Cervical length by ultrasound	0.74	0.22-2.49	NS
Presence of funneling	0.42	0.03-5.97	NS
Bishop score	0.89	0.61-1.32	NS
Birth weight	71.96	4.76-1087.84	<0.005

NS, not significant.

neling, and birth weight. Of these independent variables, cervical length and gestational age at induction, but not Bishop score, were found to significantly and independently contribute to failed labor induction. Using the likelihood of cesarean delivery for failure to progress as the outcome variable, multiple logistic regression indicated that maternal height and birth weight, but not cervical length or Bishop score, were significantly and independently associated with an increased risk of cesarean delivery for failure to progress (Table 3).

DISCUSSION

Our results show that transvaginal sonographic measurements of cervical length in nulliparous women independently predict the likelihood of entering the active phase within 24 hr of induction. On the other hand, cervical length appears to poorly predict the likelihood of cesarean delivery for failure to progress in nulliparous women undergoing labor induction. These observations strongly support the notion that in nulliparous women undergoing labor induction, preinduction cervical length plays a key role in labor progression during the latent phase, but not during the active phase of labor.

Cervical status has been established to be clearly related to the outcome and success of labor induction (19). The traditional method of predicting whether labor induction will be successful involves a preinduction examination of the favorability of the cervix as assessed by the Bishop score. However, this assessment is essentially subjective and lacks reliability. Recently several investigators have demonstrated that transvaginal sonographic measurements of the cervical length are a better predictor of successful labor induction than the Bishop score (8-10). Our results concur with this finding. Several potential explanations for this finding can be considered. First, digital examination has limitations in terms of assessing change of the internal os when the external os is closed. Therefore, the entire length of the cervix and the configuration of the internal os cannot be assessed by palpation alone. Second, the different components of the Bishop score may not have equal effects on the length of latent phase (14, 20, 21). Several investigators have reported that cervical di-

latation and effacement, which are also indexes of cervical length, are independently more closely associated with successful labor induction than was Bishop scores (11, 14, 20). Therefore, the intervention of other clinical factors involved in Bishop score calculations, such as, consistency and position of the cervix, can reduce the abilities of clinical cervical length and dilatation to predict successful labor induction. Third, Bishop scores are assigned by palpating the vaginal portion of the cervix, but dynamic changes of the uterine cervix before or during labor start from the internal os and progress to the external os (22), and therefore, Bishop scores may be limited in predicting the outcome of labor induction, especially in cases with an unfavorable score. Moreover, this rationale has been confirmed by several investigators (11, 20, 21). Fourth, Bishops scores are computed as an ordinal rather than a continuous variable, which may have reduced its quantification. And lastly, Bishop scores involve subjective evaluations with high interobserver variabilities (23).

An interesting finding of the present study is that neither cervical length nor Bishop score was found to be significantly associated with an increased risk of cesarean delivery for failure to progress. This finding is in line with the observations made by Roman et al. who demonstrated that the length of the active phase appears not to be correlated with cervical status prior to induction (24). Then, why does the cervical length significantly contribute to the prediction of failed labor induction and not to cesarean delivery for failure to progress? The labor component most affected by cervical length during labor induction is probably the latent phase. That is, women with long cervical lengths at the time of induction may experience a longer latent phase. On the contrary, if women, whose cervices have been short and started to dilate, begin labor by induction, not much time will elapse to enter the active phase of labor. In addition, because the majority of patients present with fully effaced cervices when the active phase begins, it is not surprising that arrest in such patients is not due to the initial cervical length. Indeed, cervical resistance, uterine contractions, and the forward pressure exerted

by the leading fetal head may have an influence on the progress of the latent phase of labor. However, after reaching complete effacement and dilatation of the cervix, disparity between the dimensions of the fetal head and maternal pelvis, malrotation, and ineffective uterine contraction may then be responsible for a higher risk of protraction or arrest disorder. As demonstrated by the present study and others (25-27), several factors other than initial cervical length, such as maternal stature, fetal weight, and nulliparity, significantly contribute to arrest disorder during the active phase and second stage of labor.

It is noteworthy that the presence of funneling was not associated with an increased risk of either failed labor induction or cesarean delivery for failure to progress. This finding is in line with the observation made by Yang et al. and Roman et al., who demonstrated that the presence of funneling did not have a significant effect on either the successful induction of labor or the duration of labor induction (time from the onset of active phase to delivery) (9, 21). However, conflicting findings were reported by Boozarjomehri et al., who found that latent-phase and total labor duration of induced labor were significantly associated with the presence of cervical wedging as noted by transvaginal ultrasonography (28). The reason for discrepant results between these studies is not clear but may be associated with the difficulty of assessing the presence of funneling at term, especially when the fetal head descends into the pelvic cavity or a mucus plug is present in the internal os. However, our data suggest a role of cervical length, but not of the presence of funneling, in the prediction of failed labor induction.

A potential limitation of the present study is that the choice of methods for cervical ripening and labor induction were made based on the Bishop score, i.e., if the cervix is unripe, vaginal PGE2 is used for preinduction cervical ripening. Although the choice of the induction agent may have an influence on the outcome of induction, our results remain valid because the physicians responsible for induction were unaware of sonographic findings, and because we did control for this potential confounding variable in the regression model. However, one question that remains, related to preparations used for preinduction cervical ripening, is whether local prostaglandins are needed in patients with a poor cervical score and a short sonographically measured cervical length. In a recent study of 80 patients, Bartha et al. reported that the use of transvaginal ultrasound rather than the Bishop score for the preinduction cervical assessment to choose induction agent significantly reduced the need for intracervical prostaglandin treatment without adversely affecting the success of induction or neonatal outcome (29). Further large, prospective, randomized studies are required to clarify this issue.

In conclusion, transvaginal sonographic measurements of cervical length were found to independently predict failed labor induction in nulliparous women, but the relatively poor predictive performance of this test indicates its limited use as a predictor of failed labor induction. Moreover, cervical length appears to poorly predict the likelihood of cesarean delivery for failure to progress in nulliparous women undergoing labor induction.

REFERENCES

1. Zhang J, Yancey MK, Henderson CE. *U.S. national trends in labor induction, 1989-1998. J Reprod Med 2002; 47: 120-4.*

2. Seyb ST, Berka RJ, Socol ML, Dooley SL. *Risk of cesarean delivery with elective induction of labor at term in nulliparous women. Obstet Gynecol 1999; 94: 600-7.*

3. Yeast JD, Jones A, Poskin M. *Induction of labor and the relationship to cesarean delivery: a review of 7001 consecutive inductions. Am J Obstet Gynecol 1999; 180: 628-33.*

4. Vrouenraets FP, Roumen FJ, Dehing CJ, van den Akker ES, Aarts MJ, Scheve EJ. *Bishop score and risk of cesarean delivery after induction of labor in nulliparous women. Obstet Gynecol 2005; 105: 690-7.*

5. Jackson GM, Ludmir J, Bader TJ. *The accuracy of digital examination and ultrasound in the evaluation of cervical length. Obstet Gynecol 1992; 79: 214-8.*

6. Iams JD, Paraskos J, Landon MB, Teteris JN, Johnson FF. *Cervical sonography in preterm labor. Obstet Gynecol 1994; 84: 40-6.*

7. Althuisius SM, Dekker GA, Hummel P, Bekedam DJ, van Geijn HP. *Final results of the Cervical Incompetence Prevention Randomized Cerclage Trial (CIPRACT): therapeutic cerclage with bed rest versus bed rest alone. Am J Obstet Gynecol 2001; 185: 1106-12.*

8. Ware V, Raynor BD. *Transvaginal ultrasonographic cervical measurement as a predictor of successful labor induction. Am J Obstet Gynecol 2000; 182: 1030-2.*

9. Yang SH, Roh CR, Kim JH. *Transvaginal ultrasonography for cervical assessment before induction of labor. J Ultrasound Med 2004; 23: 375-82.*

10. Rane SM, Guirgis RR, Higgins B, Nicolaides KH. *The value of ultrasound in the prediction of successful induction of labor. Ultrasound Obstet Gynecol 2004; 24: 538-49.*

11. Paterson-Brown S, Fisk NM, Edmonds DK, Rodeck CH. *Preinduction cervical assessment by Bishop's score and transvaginal ultrasound. Eur J Obstet Gynecol Reprod Biol 1991; 40: 17-23.*

12. Watson WJ, Stevens D, Welter S, Day D. *Factors predicting successful labor induction. Obstet Gynecol 1996; 88: 990-2.*

13. Gonen R, Degani S, Ron A. *Prediction of successful induction of labor: comparison of transvaginal ultrasonography and the Bishop score. Eur J Ultrasound 1998; 7: 183-7.*

14. Chandra S, Crane JM, Hutchens D, Young DC. *Transvaginal ultrasound and digital examination in predicting successful labor induction. Obstet Gynecol 2001; 98: 2-6.*

15. Cnattingius R, Hoglund B, Kieler H. *Emergency cesarean delivery in induction of labor: an evaluation of risk factors. Acta Obstet Gynecol Scand 2005; 84: 456-62.*

16. Rane SM, Pandis GK, Guirgis RR, Higgins B, Nicolaides KH. *Pre-induction sonographic measurement of cervical length in prolonged pregnancy: the effect of parity in the prediction of induction-to-deliv-*

ery interval. Ultrasound Obstet Gynecol 2003; 22: 40-4.

17. American College of Obstetricians and Gynecologists. *Dystocia and the augmentation of labor. ACOG technical bulletin no. 218. Washington, DC: American College of Obstetricians and Gynecologists, 1995.*

18. Park KH, Cho YK, Lee CM, Choi H, Kim BR, Lee HK. *Effect of preeclampsia, magnesium sulfate prophylaxis, and maternal weight on labor induction: a retrospective analysis. Gynecol Obstet Invest 2005; 61: 40-4.*

19. Bishop EH. *Pelvic scoring for elective induction. Obstet Gynecol 1964; 24: 266-8.*

20. Williams MC, Krammer J, O'Brien WF. *The value of the cervical score in predicting successful outcome of labor induction. Obstet Gynecol 1997; 90: 784-9.*

21. Roman H, Verspyck E, Vercoustre L, Degre S, Col JY, Firmin JM, Caron P, Marpeau L. *Does ultrasound examination when the cervix is unfavorable improve the prediction of failed labor induction? Ultrasound Obstet Gynecol 2004; 23: 357-62.*

22. Zilianti M, Azuaga A, Calderon F, Pages G, Mendoza G. *Monitoring the effacement of the uterine cervix by transperineal sonography: a new perspective. J Ultrasound Med 1995; 14: 719-24.*

23. Faltin-Traub EF, Boulvain M, Faltin DL, Extermann P, Irion O. *Reliability of the Bishop score before labour induction at term. Eur J*

24. Roman H, Verspyck E, Vercoustre L, Degre S, Col JY, Firmin JM, Caron P, Marpeau L. *The role of ultrasound and fetal fibronectin in predicting the length of induced labor when the cervix is unfavorable. Ultrasound Obstet Gynecol 2004; 23: 567-73.*

25. Harper DM, Johnson CA, Harper WH, Liese BS. *Prenatal predictors of cesarean section due to labor arrest. Arch Gynecol Obstet 1995; 256: 67-74.*

26. Sheiner E, Levy A, Feinstein U, Hallak M, Mazor M. *Risk factors and outcome of failure to progress during the first stage of labor: a population-based study. Acta Obstet Gynecol Scand 2002; 81: 222-6.*

27. Feinstein U, Sheiner E, Levy A, Hallak M, Mazor M. *Risk factors for arrest of descent during the second stage of labor. Int J Gynaecol Obstet 2002; 77: 7-14.*

28. Boozarjomehri F, Timor-Tritsch I, Chao CR, Fox HE. *Transvaginal ultrasonographic evaluation of the cervix before labor: presence of cervical wedging is associated with shorter duration of induced labor. Am J Obstet Gynecol 1994; 171: 1081-7.*

29. Bartha JL, Romero-Carmona R, Martinez-Del-Fresno P, Comino-Delgado R. *Bishop score and transvaginal ultrasound for preinduction cervical assessment: a randomized clinical trial. Ultrasound Obstet Gynecol 2005; 25: 155-9.*

Obstet Gynecol Reprod Biol 2004; 112: 178-81.