

# Transvaginal sonography of cervical length and Bishop score as predictors of successful induction of term labor: the effect of parity

P.C. Tan<sup>1</sup>, MRCOG; N. Vallikkannu<sup>1</sup>, M.D.; S. Suguna<sup>1</sup>, M.D.; K.F. Quek<sup>2</sup>, Ph.D.; J. Hassan<sup>1</sup>, Prof.

<sup>1</sup>Department of Obstetrics & Gynecology, <sup>2</sup>Department Social and Preventative Medicine, University of Malaya, Kuala Lumpur (Malaysia)

## Summary

**Objective:** To evaluate the predictive value for successful labor induction of transvaginal ultrasound (TVS) of cervical length according to parity. **Method:** TVS of the cervix was performed before term labor induction. Induction was considered successful if vaginal delivery was achieved within 24 hours; 231 women were available for final analysis. **Results:** Analysis of the receiver operator characteristics curve showed an optimal cut-off for cervical length of  $\leq 20$  mm for successful induction. Following multivariate logistic regression analysis, a sonographic short cervix (AOR 5.6;  $p < 0.001$ ) was an independent predictor of successful induction but not a favorable Bishop score ( $p = 0.47$ ). Among multiparas with a short cervix, positive and negative predictive values for successful induction were 98% (95% CI 90-100%) and 21% (95% CI 13%-32%) and among nulliparas, predictive values were 69% (95% CI 53%-82%) and 77% (95% CI 64%-87%) respectively. **Conclusion:** In nulliparas, cervical length can usefully predict labor induction outcome.

**Key words:** Transvaginal ultrasound; Cervical length; Bishop score; Labor induction; Parity.

## Introduction

Since its inception in 1964, the Bishop score [1] has remained the gold standard for assessing cervical favorability for labor induction [2]. In contemporary practice, induction of labor is done for about 20% of term pregnancies [3]. Transvaginal ultrasonic measurement of cervical length has been shown to be a better predictor of shorter induction to delivery interval when compared to Bishop Score by some investigators [4-7] but this is disputed by others [8-11]. A recent meta-analysis concludes that transvaginal sonography (TVS) is not demonstrably superior to the Bishop Score and calls for further research [2].

Parity is a strong predictor of successful induction following labor induction at term [12, 13]. Cervical length by TVS but not Bishop score has been shown to be an independent predictor of cesarean delivery following labor induction at term in nulliparas [14].

A prospective study was performed to compare TVS of cervical length and Bishop score in predicting labor induction outcome at term in a mixed population of nulliparas and multiparas. This information is useful for women contemplating labor induction to help manage their expectations of the induction process.

## Material and Methods

Women from 37 to 42 weeks of gestation with a singleton fetus, intact membranes and cephalic presentation were recruited when they presented to the induction bay of the delivery suite for labor induction. Recruitment was carried out by the

investigators. Women with intrauterine fetal death or known gross fetal anomaly were excluded.

Recruitment took place from January 2003 to October 2005. Two hundred and fifty-three women were approached. One woman declined TVS. Another woman with a previous cesarean scar opted for elective cesarean delivery after TVS but before commencement of labor induction. Two other women who were preterm (at 35 and 36 weeks of gestation) were mistakenly recruited. These four women were excluded. We had previously reported on tolerability of TVS versus digital assessment for Bishop score in this group of 249 women and we used all-cause cesarean delivery as the primary outcome measure [13]. Eighteen women delivered by cesarean section for non-reassuring fetal status within 24 hours of labor induction were excluded for this analysis [5] these women were excluded as having non-reassuring fetal status soon after induction of labor likely due to a pre-existing condition e.g., uteroplacental insufficiency. We felt such cases were potential confounders in the analysis of assessment methods of cervical favorability for labor induction. Following these exclusions, 231 women were left for the final analysis.

We defined successful labor induction as vaginal delivery within 24 hours of labor induction [15]. Cesarean delivery indicated by failure to progress after labor induction that took place within 24 hours of labor induction was considered as a failed induction, as failure to progress could be considered a failure of current methods of labor induction to overcome an unfavorable cervix. All other deliveries after 24 hours of commencement were also considered to be failed labor inductions.

Women were categorized into two groups based on their parity: nulliparas (no previous births) and multiparas (at least 1 previous birth). We further defined short stature as maternal height of  $< 150$  cm, post date as gestational age  $> 40$  weeks and categorized maternal age into under 35 years old and 35 years and above to assess the effect of these parameters on successful labor induction.

Study women were asked to empty their bladder before TVS.

Revised manuscript accepted for publication June 19, 2008

Ultrasound was performed in a specially equipped room separate from the induction bay. The sagittal image of the entire cervical canal was acquired sonographically as previously described [16]. Measurement of cervical length was made from the internal os to external os in a straight line [17]. The cervical images were acquired three times in succession and cervical length was measured for each image. The shortest length of the three obtained was used for analysis [5]. Funneling was defined as funnel shape appearance at the internal cervical os due to dilatation of the internal os of at least 5 mm as detected by TVS. Ultrasound findings were concealed from providers.

The ultrasound examination was performed by investigators who had at least a year's experience with sonography. Sonography was performed using a Toshiba Eccocee or Toshiba Capasee ultrasound machine (Toshiba Medical Systems, Tokyo, Japan) with a 6 MHz endovaginal probe.

The Bishop score was obtained by induction bay providers blinded to the ultrasound findings.

The method of labor induction was decided upon by providers. Labor was induced by dinoprostone (3 mg) pessary or amniotomy in our institution as previously described [13]. Universal electronic fetal monitoring was performed at labor induction.

After dinoprostone pessary insertion, a planned assessment was made after six hours and depending on cervical favorability, another dinoprostone pessary (maximum 2 doses in 24 hours) might be inserted or amniotomy performed. The women were then assessed again after six hours: if the cervix remained unfavorable and immediate delivery was not indicated, the women usually rested overnight for the induction to be repeated the following morning. If labor induction was by amniotomy, oxytocin was given within two hours depending on contractions. When labor was established, vaginal assessment was done at least every four hours.

The study was approved by our institution review board and written consent was obtained from participants. This study was done at a university hospital conducting about 5,000 deliveries per year.

Data was entered into a statistical software package SPSS version 13.0 (SPSS Inc., Chicago IL) and GraphPad Instat software (GraphPad Software Inc., San Diego CA) was also used for data analysis. The Student's *t* test was used to analyze means, Fisher's exact test was used for categorical datasets, relative risk (RR) and its 95% confidence interval (CI) were calculated using GraphPad Instat. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) (and their 95% confidence interval, CI) and likelihood ratio of a test were obtained using GraphPad Instat. A receiver operating characteristic curve was used to determine the best cut-off for cervical length and Bishop score for successful labor induction, and area under the curve was derived. Multiple logistic regression analysis to determine independent predictors of successful labor induction was performed and all variables with crude  $p < 0.05$  were included in the analysis;  $p < 0.05$  at 2-tailed was taken as a significant level.

## Results

Data from 231 women were analyzed. Labor induction was successful (vaginal delivery within 24 hours of commencement of induction) in 157/231 (68%) comprising 149 spontaneous vaginal deliveries and eight vaginal instrumental deliveries. Seventy-four women failed induction comprising 11 women who were delivered by

cesarean section for failure to progress within 24 hours and another 63 women (37 vaginal deliveries: 24 spontaneous and 13 instrumental vaginal deliveries and 26 cesarean sections: 19 for failure to progress and 7 for fetal distress) who delivered more than 24 hours after induction. Table 1 showed the characteristics and outcomes of 231 study women stratified according to vaginal delivery within 24 hours of labor induction (successful induction).

Receiver operating characteristics (ROC) curve indicated that the best cut-off for cervical length as predictor for successful labor induction was  $\leq 20$  mm (sensitivity 54%, specificity 81%, PPV 86% and NPV 45%). The optimal cut off for Bishop score was a score of  $\geq 6$  (sensitivity 57%, specificity 69%, PPV 80% and NPV 43%). Although cervical length  $\leq 20$  mm had better positive and negative predictive values for successful labor induction than Bishop score  $\geq 6$ , their 95% confidence intervals overlapped. ROC area under the curve analyses and Fisher's exact test both showed that cervical length and Bishop score were useful predictors of successful labor induction. Cervical length of  $\leq 20$  mm had better predictive values for success of induction compared to Bishop Score  $\geq 6$  (Table 2).

Table 1 also shows the interaction of sonographic findings and pre-induction clinical variables on successful labor induction. On univariate analysis parity, cervical length, funneling at the internal cervical os, maternal age, Bishop score and mode of induction were all associated with successful induction. After adjustment, controlling for the significant variables stated above, only parity  $\geq 1$  [(adjusted odds ratio (AOR) 12.4; 95% CI 5.6-27.3;  $p < 0.001$ ), cervical length  $\leq 20$  mm (AOR 5.6; 95% CI 2.3-13.8;  $p < 0.001$ ) and labor induction by amniotomy (AOR 6.5; 95% CI 1.3-32,  $p = 0.023$ ) were independently associated with successful labor inductions.

We performed a further analysis stratified by parity. Among nulliparas and multiparas, the PPVs and NPVs for successful labor induction again favored cervical length  $\leq 20$  mm compared to a Bishop score  $\geq 6$  (Table 3) though their 95% CIs overlapped. For short cervical length  $\leq 20$  mm in nulliparas, the PPVs and NPVs for successful labor induction were 69% and 77%, respectively as compared to 98% and 21% among the multiparas.

## Discussion

Our finding added to the available evidence [4-7] that TVS is better than Bishop score in predicting successful labor induction. Our study had also shown that Bishop score was no longer independently associated with successful labor induction once parity, cervical length and mode of labor induction were taken into account in a multivariate logistic regression model as has previously been demonstrated [12].

The cut-offs for cervical length of  $\leq 20$  mm and Bishop score  $\geq 6$  by ROC for predicting vaginal delivery within 24 hours determined in this analysis is identical to that

Table 1. — Characteristics. Number (%) or mean  $\pm$  standard deviation.

	Successful labor induction n = 157	Failed labor induction n = 74	Crude p value	Adjusted Odds Ratio (AOR) (95% confidence interval) and p value
Age (years)	30.5 $\pm$ 4.9	29.7 $\pm$ 4.3	p = 0.26	
$\geq 35$ years	32 (82.1)	7 (17.9)	p = 0.04	AOR 1.2 (0.4-3.4): p = 0.73
< 35 years	125 (65.1)	67 (34.9)		
Parity	1.4 $\pm$ 1.2	0.4 $\pm$ 1.0	p < 0.001	
Nulliparas	42 (42.4)	57 (57.6)	p < 0.001	AOR 12.4 (5.6-27.3): p < 0.001
Multiparas	115 (87.1)	17 (12.9)		
Ethnicity				
Malay	105 (69.5)	46 (30.5)	p = 0.57	
Chinese	29 (72.5)	11 (27.5)		
Indian	19 (52.8)	11 (47.2)		
Others	4 (100)	0 (0)		
Height (cm)	156 $\pm$ 6	156 $\pm$ 6	p = 0.84	
< 150 cm	20 (76.9)	6 (23.1)	p = 0.38	
$\geq 150$ cm	136 (67.3)	66 (32.7)		
Gestation age (weeks)	39.9 $\pm$ 1.2	39.9 $\pm$ 1.4	p = 0.77	
> 40 weeks gestation	69 (63.3)	40 (36.7)	p = 0.16	
$\leq 40$ weeks	88 (72.1)	34 (27.9)		
Indications for induction of labor†				
Diabetes mellitus	63 (74.1)	22 (25.9)	p = 0.24	
Prolonged pregnancy	54 (66.7)	27 (33.3)		
Non reassuring fetal status§	26 (66.7)	13 (33.3)		
Hypertension	19 (52.8)	17 (47.2)		
Others	16 (72.7)	6 (27.3)		
Cervical length (mm)	21 $\pm$ 9	26 $\pm$ 7	p < 0.001	
Cervix > 20 mm	73 (54.9)	60 (45.1)	p < 0.001	AOR 5.6 (2.3-13.8): p < 0.001
Cervix $\leq 20$ mm	84 (85.7)	14 (14.3)		
Funneling at internal cervical os				
No	116 (63.0)	68 (37)	p = 0.001	AOR 1.3 (0.4-4.3): p = 0.69
Yes	41 (87.2)	6 (12.8)		
Bishop score	6.0 $\pm$ 1.8	4.8 $\pm$ 1.4	p < 0.001	
Bishop score $\leq 5$	68 (57.1)	51 (42.9)	p < 0.001	AOR 1.3 (0.6-3.0): p = 0.47
Bishop score $\geq 6$	89 (79.5)	23 (20.5)		
Mode of labor induction				
Vaginal dinoprostone	104 (59.1)	72 (40.9)	p < 0.001	AOR 6.5 (1.3-32.0): p = 0.023
Amniotomy	53 (96.4)	2 (3.6)		
Induction to delivery interval (hours)	9.5 $\pm$ 5.2	34.0 $\pm$ 13.1	p < 0.001	

\*Successful labor induction defined by vaginal delivery within 24 hours of labor induction.

†Multivariate logistic regression analysis using all variables with crude p < 0.05. Adjusted odds ratio shown where variable is used in the model.

‡Total of 263 indications as 30 women had 2 indications and one had 3 indications listed for induction of labor.

§Non reassuring fetal status included oligohydramnios, suspected intrauterine growth restriction, reduced fetal movement and suboptimal umbilical artery Doppler or cardiotocograph.

Table 2. — Comparison of transvaginal sonography for cervical length  $\leq 20$  mm and Bishop score  $\geq 6$  for predicting successful labor successful induction (vaginal delivery within 24 hours) in 231 women. 95% confidence interval within brackets (95% CI).

	Prediction of vaginal delivery within 24 hours of induction of labor*	
	Cervical length $\leq 20$ mm vs > 20 mm	Bishop score Score > 5 vs Score $\leq 5$
Relative risk and p value	RR 1.6 (1.3-1.9): p < 0.001	RR 1.4 (1.2-1.7): p < 0.001
Sensitivity	54% (45%-61%)	57% (49%-65%)
Specificity	81% (70%-89%)	69% (57%-79%)
Positive predictive value	86% (77%-92%)	80% (71%-86%)
Negative predictive value	45% (37%-54%)	43% (34%-52%)
Likelihood ratio	2.8	1.8
Area under receiver operator characteristics curve	0.68 (0.61-0.75): p < 0.001	0.70 (0.63-0.77): p < 0.001

\*Eighteen cesarean deliveries indicated by fetal distress within 24 hours of labor induction censored and excluded from analysis.

reported in our previous report on all-cause cesarean delivery as outcome which has also found sonographic cervical length but not Bishop score to be an independent predictor [13]. This finding suggests that cervical length is a reliable independent predictor over a range of definitions of failed inductions with consistent cervical length cut-offs.

In this analysis, funneling at the internal cervical os on ultrasonic assessment was not an independent predictor of cesarean delivery in contrast to an earlier finding [6] due to the strong and dependent association between funneling and short cervical length (p < 0.001) within our study.

Multiparity [12, 18] and mode of induction [1, 10] as anticipated were significant predictors for successful labor induction after adjusted analysis.

Our data indicated that there was a distinct difference in the interpretation of cervical length as a predictor of

Table 3. — Comparison of transvaginal sonography for cervical length  $\leq 20$  mm and Bishop score  $\geq 6$  for successful labor induction (vaginal delivery within 24 hours) in labor induction at term stratified according to parity in 231 study women. 95% confidence interval within brackets (95% CI).

	Prediction of successful induction: Nulliparas		Prediction of successful induction: Multiparas	
	Cervical length $\leq 20$ mm v $> 20$ mm	Bishop score Score $> 5$ v Score $\leq 5$	Cervical length $\leq 20$ mm v $> 20$ mm	Bishop score Score $> 5$ v Score $\leq 5$
Relative risk and p value	RR 3.0 (1.8-5.1): p < 0.001	RR 1.9 (1.2-3.0): p = 0.014	RR 1.2 (1.1-1.4): p = 0.001	RR 1.2 (1.1-1.4): p = 0.008
Sensitivity	69% (53%-82%)	62% (46%-76%)	47% (38%-57%)	55% (45%-64%)
Specificity	77% (64%-87%)	65% (51%-77%)	94% (71%-100%)	82% (57%-96%)
Positive predictive value	69% (53%-82%)	56% (41%-71%)	98% (90%-100%)	95% (87%-99%)
Negative predictive value	77% (64%-87%)	69% (56%-82%)	21% (13%-32%)	21% (12%-33%)
Likelihood ratio	3.0	1.8	8.1	3.1

\*Eighteen women who had emergency cesarean sections for fetal distress within 24 hours of labor induction censored and excluded leaving 231 women for final analysis.

successful labor induction at term between multiparas and nulliparas. The PPV of short cervical length  $\leq 20$  mm for successful labor induction of 98% (95% CI, 90-100%) in multiparas was very high: 98% of these women might deliver vaginally within 24 hours. However, in contrast, the NPV was only 21% (95% CI, 12-33%) which was very poor: even with long cervical length  $> 20$  mm, 79% were expected to deliver vaginally within 24 hours. Among nulliparas, short cervical length  $\leq 20$  mm had a PPV of 69% (95% CI 53-82%) and a NPV of 77% (64-87%). This implied that 69% of nulliparous women with a short cervix might expect vaginal delivery within 24 hours of labor induction and 77% of nulliparas with a long cervix might encounter failure to deliver vaginally within 24 hours.

A previous study has shown TVS cervical length  $< 27$  mm to have a sensitivity of 76% and specificity of 75.5% for predicting successful labor induction and vaginal delivery in nulliparas [14]. Our sensitivity and specificity at 69% and 77%, respectively, for cervical length  $\leq 20$  mm in predicting successful labor induction among nulliparas were comparable to the earlier study.

A potential problem with the use of TVS cervical length in assessing cervical favorability for labor induction is the different cut-offs suggested by previous studies ranging from 24 mm to 30 mm for the cervical length [4, 5, 7, 14]. Our study cut-off at 20 mm is similar to 18-20 mm cut-off of a recent large study [19].

Interestingly, a recent randomized study comparing TVS assessment and Bishop score to guide pre-induction cervical ripening with prostaglandins has shown a reduction in prostaglandin use with TVS without affecting the success rate of labor induction [20].

Previous studies have also shown that TVS is better tolerated compared to digital assessment for Bishop score [10, 13, 21].

## Conclusion

TVS for cervical length may be a better alternative to assess favorability for labor induction for nulliparous women in the office setting provided expertise and ultrasound equipment are available. Multiparas respond well to labor induction regardless of cervical favorability.

## References

- [1] Bishop E.H.: "Pelvic scoring for elective induction". *Obstet. Gynecol.*, 1964, 24, 269.
- [2] Crane J.M.: "Transvaginal ultrasound cervical length and successful labor induction: a systematic review". *Obstet. Gynecol.*, 2006, 107 (suppl. 4), 60S.
- [3] Government Statistical Service. NHS Maternity Statistics, England: 2003-04. Bulletin 2005/10. March 2005. Accessible at <http://www.dh.gov.uk/assetRoot/04/10/70/61/04107061.pdf>.
- [4] Yang S.H., Roh C.R., Kim J.H.: "Transvaginal ultrasonography for cervical assessment before induction of labor". *J. Ultrasound Med.*, 2004, 23, 375.
- [5] Pandis G.K., Papageorgiou A.T., Ramanathan V.G., Thompson M.O., Nicolaides K.H.: "Preinduction sonographic measurement of cervical length in the prediction of successful induction of labor". *Ultrasound Obstet. Gynecol.*, 2001, 18, 623.
- [6] Boozarjomehri F., Timor-Tritsch I., Chao C.R., Fox H.E.: "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] Gomez Laencina A.M., Sanchez F.G., Gimenez J.H., Martinez M.S., Valverde Martinez J.A., Vizcaino V.M.: "Comparison of ultrasonographic cervical length and the Bishop score in predicting successful labor induction". *Acta Obstet. Gynecol. Scand.*, 2007, 86, 799.
- [8] Rozenberg P., Chevret S., Chastang C., Ville Y.: "Comparison of digital and ultrasonographic examination of the cervix in predicting time interval from induction to delivery in women with a low Bishop score". *BJOG*, 2005, 112, 192.
- [9] Reis F.M., Gervasi M.T., Florio P., Bracalente G., Fadalti M., Severi F.M. *et al.*: "Prediction of successful induction of labor at term: role of clinical history, digital examination, ultrasound assessment of the cervix, and fetal fibronectin assay". *Am. J. Obstet. Gynecol.*, 2003, 189, 1361.
- [10] Chandra S., Crane J.M., Hutchens D., Young D.C.: "Transvaginal ultrasound and digital examination in predicting successful labor induction". *Obstet. Gynecol.*, 2001, 98, 2.
- [11] 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.
- [12] Peregrine E., O'Brien P., Omar R., Jauniaux E.: "Clinical and ultrasound parameters to predict the risk of cesarean delivery after induction of labor". *Obstet. Gynecol.*, 2006, 107, 227.
- [13] Tan P.C., Vallikkannu N., Suguna S., Quek K.F., Hassan J.: "Transvaginal sonographic measurement of cervical length vs. Bishop score in labor induction at term: tolerability and prediction of cesarean delivery". *Ultrasound Obstet. Gynecol.*, 2007, 29, 568.
- [14] Daskalakis G., Thomakos N., Hatzioannou L., Mesogitis S., Papantoniou N., Antsaklis A.: "Sonographic cervical length measurement before labor induction in term nulliparous women". *Fetal Diagn. Ther.*, 2006, 21, 34.
- [15] Rane S.M., Guirgis R.R., Higgins B., Nicolaides K.H.: "Models for the prediction of successful induction of labor based on pre-induction sonographic measurement of cervical length". *J. Matern. Fetal. Neonatal Med.*, 2005, 17, 315.



- [16] Burger M., Weber-Rossler T., Willmann M.: "Measurement of the pregnant cervix by transvaginal sonography: an interobserver study and new standards to improve the interobserver variability". *Ultrasound Obstet. Gynecol.*, 1997, 9, 188.
- [17] To M.S., Skentou C., Chan C., Zagaliki A., Nicolaides K.H.: "Cervical assessment at the routine 23-week scan: standardizing techniques". *Ultrasound Obstet. Gynecol.*, 2001, 17, 217.
- [18] Rane S.M., Guirgis R.R., Higgins B., Nicolaides K.H.: "The value of ultrasound in the prediction of successful induction of labor". *Ultrasound Obstet. Gynecol.*, 2004, 24, 538.
- [19] Rane S.M., Guirgis R.R., Higgins B., Nicolaides K.H.: "Pre-induction sonographic measurement of cervical length in prolonged pregnancy: the effect of parity in the prediction of the need for cesarean section". *Ultrasound Obstet. Gynecol.*, 2003, 22, 45.
- [20] Bartha J.L., 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.
- [21] Paterson-Brown S., Fisk N.M., Edmonds D.K., Rodeck C.H.: "Preinduction cervical assessment by Bishop's score and transvaginal ultrasound". *Eur. J. Obstet. Gynecol. Reprod. Biol.*, 1991, 40, 17.

Address reprint requests to:  
P.C. TAN, M.D.  
Department of Obstetrics & Gynecology,  
Faculty of Medicine, University of Malaya,  
Lembah Pantai,  
50603 Kuala Lumpur (Malaysia)  
e-mail: pctan@um.edu.my