Bishop score and three-dimensional ultrasound assessment of the cervix for predicting time to spontaneous onset of labor and time to delivery in post-term pregnancy

L. Rovas[1], MD, P. Sladkevicius[1], MD, PhD, E. Strobel[1], RN, RM, F. De Smet[2], MD, MscEng, PhD, B. De Moor[2], MscEng, PhD, L. Valentin[1], MD, PhD

[1] Obstetric, gynecological and prenatal ultrasound research unit, Department of Clinical Sciences, Malmö, Lund University, Sweden.

[2] Department of Electrical Engineering, ESAT-SCD, Katholieke Universiteit Leuven, Leuven, Belgium.

Department of Obstetrics and Gynecology, Malmo University Hospital, Malmo, Sweden.

Short title: 3D ultrasound and cervix.

Key words: cervix uteri, three-dimensional imaging, Doppler ultrasound, spontaneous delivery, prolonged pregnancy.

Correspondence: Povilas Sladkevicius, Kvinnokliniken, Universitetssjukhuset MAS, S-20502 Malmo, Sweden.

Telephone: +46 40 332094 Fax: +46 40 962600

Email: Povilas.Sladkevicius@obst.mas.lu.se

ABSTRACT

Objectives To determine if three-dimensional (3D) power Doppler ultrasound examination of the cervix is useful to predict time to spontaneous **onset** of labor or time to delivery in post-term pregnancy.

Methods A prospective study was conducted in 60 women who went into spontaneous labor. All underwent transvaginal 3D power Doppler ultrasound examination of the cervix immediately before a planned routine post-term check-up at ≥ 41 gestational weeks + 5 days. The variables analyzed were length, anterior-posterior (AP) diameter and width of the cervix and of any cervical funneling, cervical volume (cm^3), vascularization index (VI), flow index (FI) and vascularization flow index (VFI), parity and Bishop score. Univariate and multivariate logistic regression were used to determine which variables predicted spontaneous **onset** of labor >24h and >48h and of vaginal delivery > 48h and > 60h. **Multivariate logistic regression analysis was done both with and without including Bishop score as predictive variable.**

Results The areas under ROC curves did not differ significantly between the diagnostic tests (Bishop score, cervical length, logistic regression models)

Time to onset of labor

The likelihood of **onset** of labor > 24h increased with increasing cervical length and width. The area under the ROC curve for a logistic regression model including cervical width and length was 0.839, the best cutoff for likelihood was 0.56,and the positive (LR+) and negative (LR-) likelihood ratios to predict **onset** of labor > 24h were 4.8 and 0.21,respectively. The likelihood of **onset o**f labor > 48 h was lower in parous women and increased with increasing cervical length. The area under the ROC curve for a logistic regression modelincluding parity and cervical length was 0.788, the best cutoff for likelihood was 0.44, and the LR+ and LR- to predict onset of labor > 48 h were 3.0 and 0.38, respectively.

Time to vaginal delivery

When Bishop score was included in the regression model, it was the only variable independently related to vaginal delivery >48 h, the likelihood of delivery > 48 h increasing with decreasing Bishop score (area under the ROC curve 0.816, best cutoff for Bishop score 5, LR+ and LR- to predict delivery > 48 h 3.1 and 0.21). If Bishop score was not included in the logistic regression model, cervical length, VI and FI independently predicted delivery > 48h, the likelihood of delivery > 48h increasing with increasing cervical length, decreasing VI and increasing FI (area under ROC curve for the logistic regression model 0.805, best cut-off for likelihood 0.39, LR+ and LR- 3.7 and 0.19). Cervical length and Bishop score were both independent predictors of delivery > 60h (area under ROC curve for the logistic regression model 0.806, best cutoff for likelihood 0.38, LR+ and LR– 3.3 and 0.22).

Conclusion. Bishop score, sonographic cervical length, and various logistic regression models predict time to spontaneous onset of labor and time to vaginal delivery with similar accuracy in post-term pregnancy. Results of 3D power Doppler ultrasound examination are related to time to delivery >48h.

Introduction

Efforts have been made to understand the cervical ripening and opening process, but we still have insufficient knowledge about why some women deliver pre-term and others post-term. The traditional method to investigate cervical readiness for labor is digital examination, where the results of the examination are summarized in a Bishop score[1]. Although the Bishop score is recognized as a useful method, there are problems with its accuracy, because half of the cervix is not palpable at vaginal examination if the cervical canal is closed[2]. Transvaginal ultrasound allows visualization of the entire cervix irrespective of whether the internal cervical os is closed or open at palpation. Studies have shown transvaginal ultrasound examination of the cervix to be a predictor of preterm, term or post-term delivery, and outcome of labor induction[3-6]. However, one research team found no relationship between sonographically measured cervical length and time to spontaneous **onset** of labor[7]. A few studies have suggested three-dimensional (3D) ultrasound examination of the cervix to allow a more complete assessment of the cervix than two-dimensional (2D) ultrasound examination[8,9].

There is evidence that angiogenic factors may play a role in cervical ripening and the birth process[10]. Therefore, we speculated that there might be changes in cervical vascularization during cervical ripening potentially detectable by 3D power Doppler ultrasound examination, and that therefore 3D power Doppler ultrasound examination might be useful to predict time to spontaneous **onset** of labor.

The aim of this study was to determine if 3D ultrasound examination of the cervix including 3D power Doppler ultrasound examination can provide useful information to predict the time to spontaneous **onset** of labor or to vaginal delivery in **post-term** pregnancy.

Material and methods

The Ethics Committee of the Medical faculty of Lund University, Sweden approved the study protocol. Informed written consent was obtained from all women, after the nature of the procedures had been fully explained to them.

Pregnant women are routinely seen in our antenatal outpatient department for a routine post-term check-up of the mother-to-be and fetus at 42 gestational weeks (gws) and 0 days (42 + 0 gws). If a pregnant woman reaches 42 + 0 gws during a weekend, the assessment is done on the preceding weekday. As a result of this, some women are examined at gws 41 + 5 or 41 + 6. Based on the results of the examinations of the fetus and mother, the managing obstetrician chooses either to induce labor or to await spontaneous **onset** of labor. Post-term check-ups are scheduled to take place at 7.30 a.m. All pregnant women coming to our antenatal outpatient department for their first post-term check-up were asked to participate in our study, i.e., to undergo a transvaginal 3D power Doppler ultrasound examination of the cervix immediately before their ordinary routine post-term check-up.

Inclusion criteria for this study were: singleton pregnancy at ≥ 41 + 5 gws, gestational age determined by ultrasound fetometry at 14 – 20 gws, live fetus, vertex presentation, spontaneous **onset** of labor, no previous cone biopsy, intact membranes, no bleeding and not in labor at ultrasound examination, no digital examination of the cervix ≤ 24 h before the ultrasound examination.

Of 121 consecutive women asked to participate, six women declined participation. Thus, a transvaginal ultrasound examination of the cervix was performed in 115 women. Sixty of these women fulfilled all our inclusion criteria and constitute our study population.

Transvaginal sonography was carried out by the first author as described below. The equipment used was a GE Voluson 730 Expert ultrasound system (General Electrics, Zipf, Austria) equipped with a 2.8 – 10 MHz transvaginal transducer. The field of view was 146°. Identical pre-installed ultrasound settings were used in all women. The power Doppler settings were: frequency 3 – 9 MHz, pulse repetition frequency 0.6 kHz, gain -5.0, wall motion filter 'low 1'. The women were examined in the lithotomy position with an empty bladder. The ultrasound probe was slowly introduced into the vagina and care was taken to avoid exerting undue pressure, which may artificially lengthen the cervix[11]. After a satisfactory image had been obtained, the probe was withdrawn until the image became blurred. Then the probe was gradually advanced with only enough pressure to restore a satisfactory image. A sagittal plane through the cervix was selected where the internal os, the cervical canal and the external os were visible simultaneously. Neither fundal nor suprapubic pressure was applied. After obtaining a good 2D gray scale ultrasound image of the cervix, the system was switched into the 3D mode, and then into the power Doppler mode. A longitudinal section of the cervix was centralized within the 3D sector appearing on the ultrasound screen. An ultrasound volume, containing the cervix, was acquired by holding the transducer stationary while its crystals were mechanically rotated across the sector with a sweep angle of 90°. The duration of the volume acquisition was 15 – 20 seconds depending on the dimensions of the 3D sector. The acquired volumes were stored on the hard disk of the ultrasound system for later analysis off-line. The following measurements were taken using 'any-plane' slicing of the volume acquired: length, anterior-posterior (AP) diameter and width of the cervix and of any cervical funneling (Figure 1). Funneling was defined as any visible opening of the internal cervical os. Cervical volume (cm^3) and power Doppler flow indices were calculated using the virtual organ computer aided analysis software

(VOCAL™), which is integrated into the GE Voluson 730 Expert ultrasound system. The following Doppler indices were calculated as described earlier[12,13]: vascularization index (VI), flow index (FI) and vascularization flow index (VFI); see Figure 1. All ultrasound results were unavailable to the clinical staff.

After the ultrasound examination a digital examination of the cervix (Bishop score) was performed by the obstetrician in charge in the labor ward. The results were noted in a dedicated study form. **Onset** of labor was defined as the time when uterine contractions were regular with at least two contractions per 10 minutes. Clinical information about the patients was obtained from their medical records.

Statistical calculations were made using the Statistical Package for the Social Sciences (SPSS Inc., Chicago, Illinois, USA, version 12.02). The following outcome variables were chosen à priori: **onset** of labor > 24h and > 48h, and delivery > 48h, and > 60h after the ultrasound examination. The statistical significance of a possible relationship between the outcome and the background variables was determined using univariate logistic regression with likelihood ratios, the background variables examined being parity (multipara vs. nullipara), gestational age (< 42 + 0 gws vs. ≥ 42 + 0 gws), Bishop score, cervical length, AP diameter and width, flow indices, funneling (yes vs. no), and funnel size. Multivariable logistic regression with stepwise selection of variables was used to determine which variables were independently associated with the outcome. The multivariable logistic regression analyses were made including parity (nullipara coded as 0 vs. multipara coded as 1), gestational age (< 42 + 0 gws coded as 0 vs. ≥ 42 + 0 gws coded as 1), and results of 3D ultrasound examination with and without Bishop score as predicting variables (all measurements in mm, funneling expressed as the mean of length, AP diameter and width with 0 indicating absence of funneling). The likelihood ratio test was used to determine which variables to include in the logistic regression model, a P-value < 0.05 being the threshold for inclusion. The objective of the model building process was to obtain a 'good fit' for the data with the least number of independent variables. The application of the regression equation to data from each woman gave the probability for that woman to go into spontaneous labor > 24h or > 48h, and to deliver > 48h or > 60h, the estimated probability ranging from 0 to 1. Receiver operating characteristic (ROC) curves[14] were drawn for each diagnostic test (i.e., Bishop score, cervical length, and the logistic regression models) to evaluate its diagnostic ability. The area under the ROC curve and the 95% confidence interval (CI) of this area were calculated. If the lower limit of the CI for the area under the ROC curve was > 0.5, the diagnostic test was considered to have a discriminatory potential. The statistical significance of a difference in the area under the ROC curve between the different diagnostic tests was determined as described by Hanley and MacNeil[15,16] using a customized computer program written in MATLAB (Version 6.5.0.180913a Release 13) designed by one of the co-authors (FdS). The ROC curves were also used to determine the mathematically best cut-off value for each diagnostic test, the best cut-off value being defined as the one corresponding to the point on the ROC curve situated most far away from the reference line[14]. The sensitivity, false positive-rate, and positive and negative likelihood ratio of the optimal cut-off values with regard to predicting **onset** of labor > 24h, > 48h, delivery > 48h and > 60h were also calculated. Two tailed p-values ≤ 0.05 are considered statistically significant.

Results

Mean age was 31 years ± 4.9 (SD; standard deviation) and mean body mass index in the first trimester 28.7 kg/m^2 ± 3.09. Twenty-nine (48%) women were parous. Nineteen women were examined at 41 + 5 – 6 gws, 37 women at 42 + 0 gws, and four women at 42 + 1 – 4 gws. Vaginal delivery occurred in 56 (93%) women, whereas Caesarean section was performed in four women, the indication being fetal distress (n = 2), placental abruption (n = 1), and mother's own request (n = 1). In 14 patients information about Bishop score was unavailable.

Results of univariate analyses are shown in Tables 1, 2, 3, and 4. Bishop score was lower, the cervix as measured by ultrasound was longer and cervical volume was larger in women who went into spontaneous labor > 24h and > 48h than in those who went into labor earlier, and Bishop score was lower and the cervix was longer in women who delivered > 48h and > 60h than in those who delivered earlier. VI was lower in women who went into labor > 48h and in those who delivered vaginally > 60h than in those who went into spontaneous labor/delivered earlier, and the proportion of nullipara was higher among women who delivered > 48h and > 60h, but these differences reached only marginal statistical significance (p-values 0.050 – 0.075).

Results of multivariable logistic regression analyses are shown in Table 5. Areas under ROC curves for Bishop score, cervical length, and logistic regression models, and optimal cutoff values with regard to predicting **onset** of labor > 24h, > 48h, and vaginal delivery > 48h and > 60h, and the sensitivity, false-positive rate (1 minus specificity), and positive and negative likelihood ratios for the mathematically optimal cutoff of each predicting variable are shown in Table 6. Bishop score and vascular indices did not enter any logistic regression model to predict **onset** of labor, only cervical length and width and parity did, the likelihood of **onset** of labor > 24h increasing with increasing cervical length and width and the likelihood of **onset** of labor > 48h being lower in multiparous women and increasing with increasing cervical length. If Bishop score was included in the logistic regression model building process it was the only variable independently related to vaginal delivery > 48h, the likelihood of delivery > 48h increasing with decreasing Bishop score. If Bishop score was not included in the logistic regression model, then cervical length, VI and FI independently predicted delivery > 48h, the likelihood of delivery > 48h increasing with increasing cervical length, decreasing VI and increasing FI. Cervical length and Bishop score were both independent predictors of delivery > 60h. The areas under the ROC curves did not differ significantly between the diagnostic tests (p = 0.14 – 0.92).

Discussion

Our results show that both Bishop score and ultrasonographic cervical length can predict time to spontaneous onset of labor and time to vaginal delivery in post-term pregnancy. This is in agreement with results of other studies examining which factors are associated with time to spontaneous onset of labor and/or time to spontaneous delivery[4,5,7]. Moreover, VI, which reflects the vascularization of the cervix, seems to be related to the time to vaginal delivery in post-term pregnancy, VI being higher in women giving birth < 60 h than in those giving birth later (p = 0.05).

Multivariate logistic regression showed that ultrasound variables (cervical length, cervical width) were independent predictors of time to onset of labor (> 24h and > 48h), and that Bishop score did not contribute any additional predictive information to cervical length and width. On the other hand, Bishop score was the only variable to independently predict vaginal delivery > 48h, and both Bishop score and cervical length independently predicted vaginal delivery > 60h. If Bishop score was not included as a variable in multivariate logistic regression analysis, then cervical length, VI and FI were independent predictors of delivery > 48h, whereas cervical length was the only independent predictor of delivery > 60h. A possible interpretation of our finding that increasing VI and decreasing FI increased the likelihood of delivery < 48h is that with progression of cervical ripening the density of small vessels in the cervix increases (increased VI reflects increased vessel density,FI should be low in small vessels, because small vessels contain few blood corpuscles, and therefore the back scattered energy should be low). Our findings support the theory that the cervix becomes increasingly vascularized during the cervical ripening process. To the best of our knowledge there are no other studies examining cervical vascularization in pregnancy using Doppler ultrasound techniques.

None of the variables tested with regard to their ability to predict the time to spontaneous onset of labor or to vaginal delivery were particularly good predictors. This was true both of single variables and of the logistic regression models including more than one variable. Most predictors were associated with only a small or at most a moderate change in the likelihood of the outcome, the positive likelihood ratio of the selected cut-offs never exceeding 5.8, the lowest negative likelihood ratio being 0.09[17], and the largest area under the ROC curve being 0.839 (range in areas under the ROC curve, 0.717 – 0.839). This is in agreement with results of other studies examining the ability of sonographic cervical length and Bishop score at term to predict time to onset of labor[4,5,7].

With one exception the areas under the ROC curve of our logistic regression models were larger than those of Bishop score alone or cervical length alone, suggesting the logistic regression models to be better predictors of the outcome than Bishop score alone or cervical length alone. However, the differences in area were not statistically significant, and in most cases the areas were indeed quite similar. This suggests that the methods – Bishop score, cervical length, logistic regression models – can be used interchangeably. Palpation of the cervix is cheaper than an ultrasound examination, because no technical equipment is needed, but on the other hand it is a clinical experience that it is more painful to the woman than a transvaginal ultrasound examination. Is the possibility of slightly improved prediction of onset of labor > 24h by using 3D ultrasound (cervical length and width) instead of Bishop score worth the effort? Is the possibly slightly improved prediction of onset of labor > 24h by measuring both cervical length and width and entering them into a mathematical model to calculate an individual likelihood instead of only measuring cervical length worth the extra effort? Similar questions may be asked with regard to predicting onset of labor/delivery > 48h and > 60h. Even though our results suggest that vascularization of the cervix is in some way related to cervical ripening, 3D power Doppler ultrasound examination of the cervix to predict time to vaginal delivery in post-term pregnancy is probably not a very useful clinical method because of its complexity. We can obtain similar information by the Bishop score alone. Whether the improvement in prediction of vaginal delivery > 60h by adding ultrasound examination to Bishop score is substantial enough to justify the added use of ultrasound examination is questionable. Our prediction models need to be tested prospectively.

Acknowledgments

This study was supported by the Swedish Medical Research Council (grants nos. K2001-72X-11605-06A and K2002-72X-11605-07B), two governmental grants (Landstingsfinansierad regional forskning, Region Skåne, and ALF-medel), and Funds administered by Malmö University Hospital.

References

1. Bishop EH. Pelvic scoring for elective induction. *Obstet Gynecol* 1964; **24**: 266-268.

2. Jackson GM, Ludmir J, Bader TJ. The accuracy of digital examination and ultrasound in the evaluation of cervical length. *Obstet Gynecol* 1992; **79**: 266-268.

3. Heath VCF, Southall TR, Souka AP, Elisseou A, Nicolaides KH. Cervical length at 23 weeks of gestation: prediction of spontaneous preterm delivery. *Ultrasound Obstet Gynecol* 1998; **12**: 312-317.

4. Rozenberg P, Goffinet F, Hessabi M. Comparison of the Bishop score, ultrasonographically measured cervical length, and fibronectin assay in predicting time until delivery and type of delivery at term. *Am J Obstet Gynecol* 2000; **182**: 108-113.

5. Ramanathan G, Yu C, Osei E, Nicolaides KH. Ultrasound examination at 37 weeks` gestation in the prediction of pregnancy outcome: the value of cervical assessment. *Ultrasound Obstet Gynecol* 2003; **22**: 598-603.

6. Rane SM, Guirgis RR, Higgins B, Nicolaides KH. The value of ultrasound in the predicton of successful induction of labor. *Ultrasound Obstet Gynecol* 2004; **24**: 538-549.

7. Vimercati A, Greco P, Lopalco P, Loizzi V, Scioscia M, Mei L, Rossi Selvaggi L. The value of ultrasonographic examination of the cervix in predicting post-term pregnancy. *J Perinatal Med* 2001; **29**: 317-321.

8. Severi FM, Bocchi C, Florio P, Picciolini E, D`Aniello G, Petraglia F. Comparison of two-dimensional and three-dimensional ultrasound in the assessment of the cervix to predict preterm delivery. *Ultrasound Med Biol* 2003; **29**: 1261-1265.

9. Bega G, Lev-Toaff A, Kuhlman K, Berghella V, Parker L, Goldberg B, Wapner R. Three-dimensional multiplanar transvaginal ultrasound of the cervix in pregnancy. *Ultrasound Obstet Gynecol* 2000; **16**: 351-358.

10. Mowa CN, Jesmin S, Sakuma I, Usip S, Togashi H, Yoshioka M, Hattori Y, Papka R. Characterization of vascular endothelial growth factor (VEGF) in the uterine cervix over pregnancy: effects of denervation and implications for cervical ripening. *J Histochem Cytochem*. 2004; 52: 1665-1674.

11. To MC, Skentou C, Chan C, Zagaliki A, Nicolaides HK. Cervical assessment at the routine 23-week scan: standardizing techniques. *Ultrasound Obstet Gynecol* 2001; 17: 217-219.

12. Rovas L, Sladkevicius P, Strobel E, Valentin L. Intra-observer and inter observer reproducibility of three-dimensional gray scale and power Doppler ultrasound examinations of the cervix in pregnant women. *Ultrasound Obstet Gynecol* 2005; In press.

13. Pairleitner H, Steiner H, Hasenoehrl G, Staudach A. Three-dimensional power Doppler sonography: imaging and quantifying blood flow and vascularization. *Ultrasound Obstet Gynecol*. 1999; 14:139-143.

14. Richardson DK, Schwartz JS, Weinbaum PJ, Gabbe SG. Diagnostic tests in obstetrics: a method for improved evaluation. *Am J Obstet Gynecol*. 1985; 152: 613-618.

15. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982; 143: 29-36.

16. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983; 148: 839-843.

17. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical

literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994; 271: 703-707.

Legend

Figure 1 3D ultrasound measurement of cervical volume, cervical anterior-posterior diameter, cervical width and blood flow indices. a) Multiplanar display of the cervix: longitudinal view in the upper left quadrant (used for measurement of cervical length - A and anterior-posterior diameter – B), transverse view in the upper right quadrant (used for measurement of cervical width – C), and coronal view in the lower left quadrant. The resultant 3D model is seen in the lower right quadrant. The tracing of the cervix is demarcated by lines. b) Vascular indices as shown on the ultrasound screen.

Figure 1a

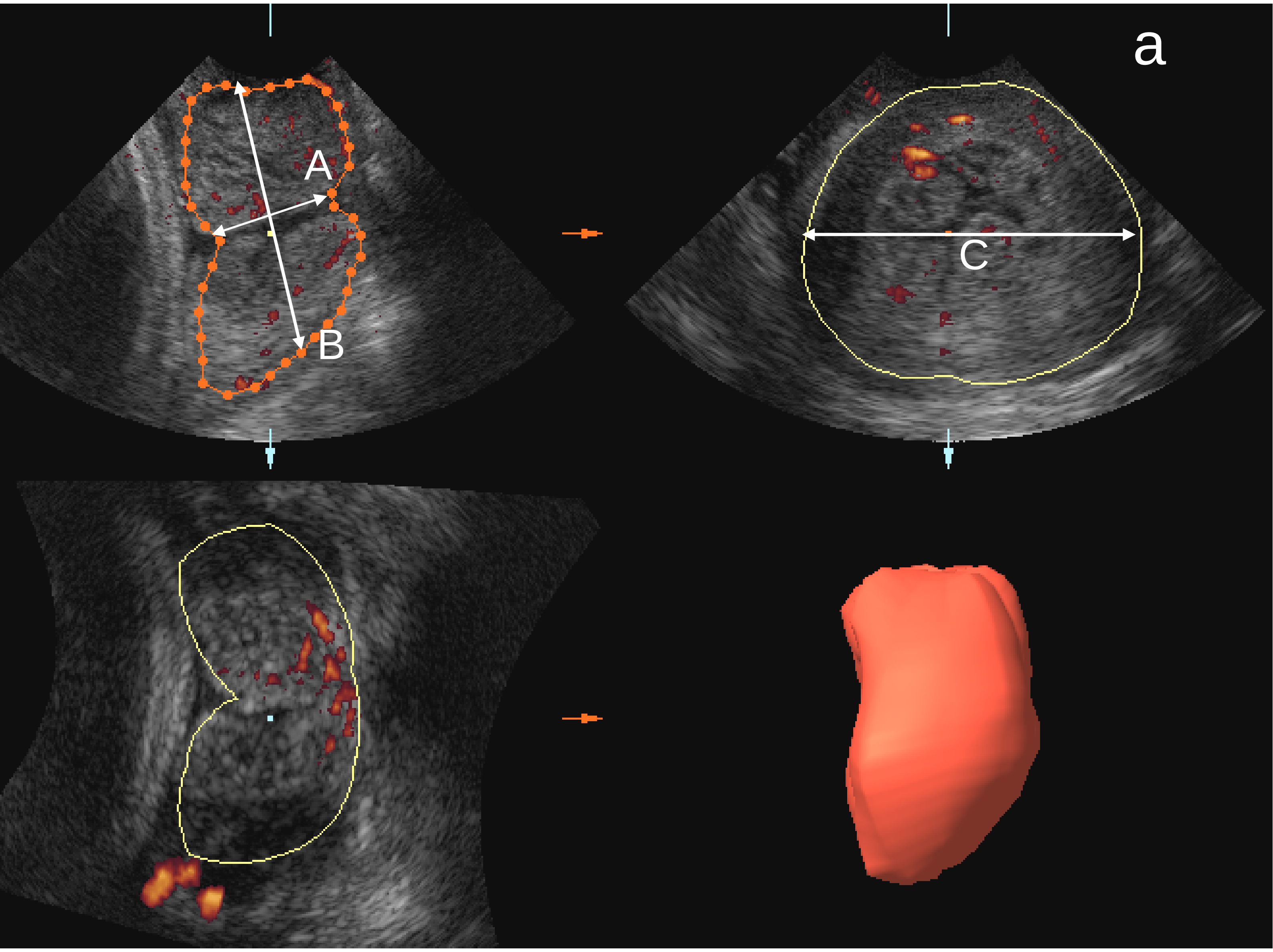

Figure 1b

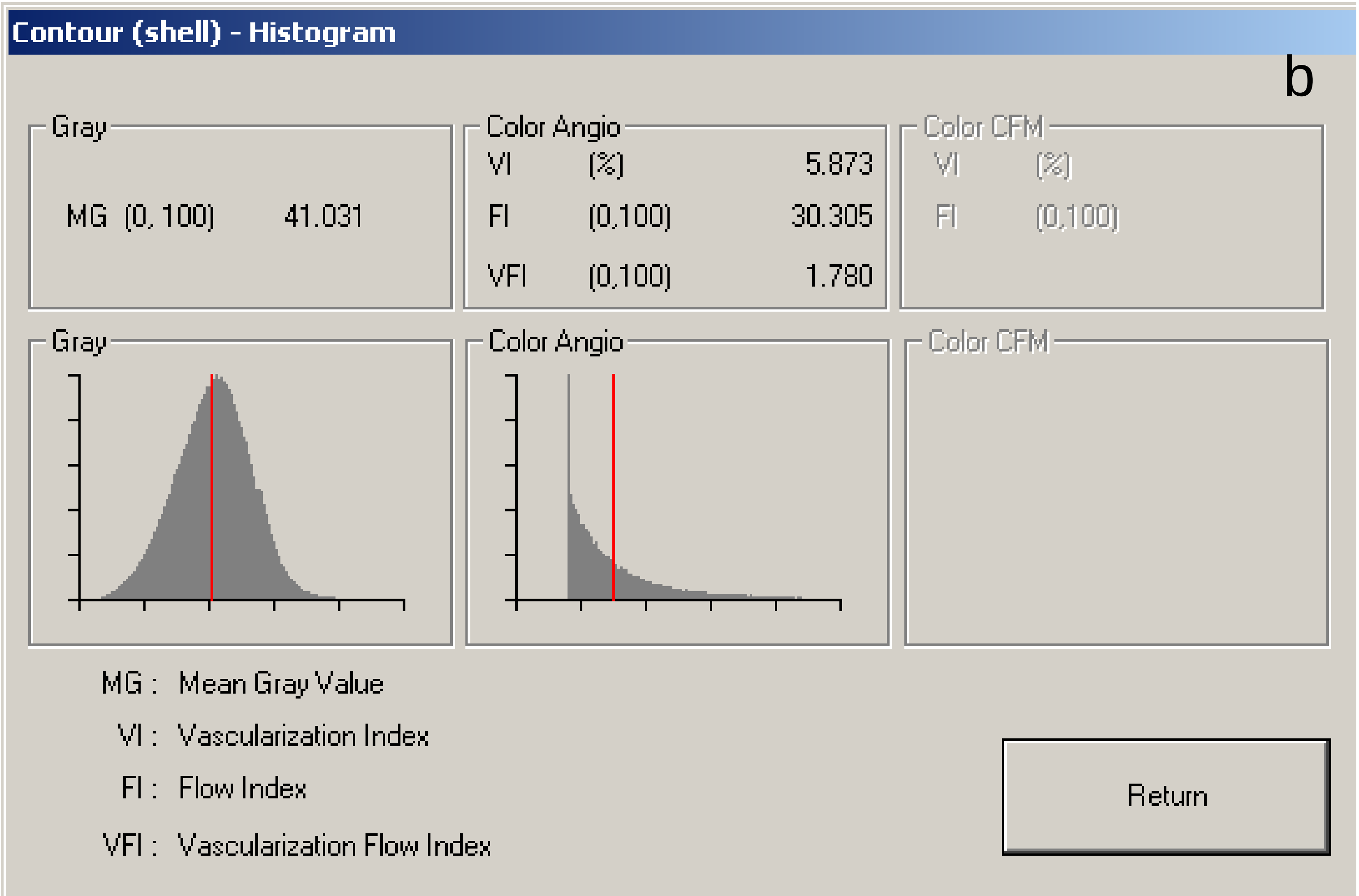
Contour (shell) - Histogram
b
Gray
MG (0, 100) 41.031
Color Angio
VI (%) 5.873
FI (0,100) 30.305
VFI (0,100) 1.780
Color CFM
VI (%)
FI (0,100)
Gray
Color Angio
Color CFM
MG : Mean Gray Value
VI : Vascularization Index
FI : Flow Index
VFI : Vascularization Flow Index
Return

Table 1. Gestational age, parity, ultrasound results, and Bishop score for women with spontaneous **onset** of labor ≤ 24 h vs. > 24 h after ultrasound examination

Results	Delivery ≤ 24 h	Delivery > 24 h	P-value
	n =21	n = 39	
Multipara; n (%)	9 (43%)	20 (51%)	0.533
< 42+0 gws at examination	8 (38%)	11 (28%)	0.435
3D ultrasound			
Cervix			
Length, mm; median (range)	7 (2 - 37)	15 (2 - 40)	0.002
AP diameter, mm; mean ± SD	38 ± 10.1	37 ± 8.5	0.874
Width, mm; mean ± SD	41 ± 5.2	44 ± 5.8	0.047
Volume, cm^3; mean ± SD	24.2 ± 10.09	33.7 ± 13.52	0.005
Flow indices			
VI; median (range)	7.5 (0.3 - 23.7)	4.5 (0.9 - 17.0)	0.268
FI; mean ± SD	29.2 ± 3.47	29.3 ± 3.98	0.891
VFI; median (range)	2.0 (0.1 - 8.3)	1.3 (0.2 - 5.6)	0.324
Funnel; n (%)	12 (57%)	24 (62%)	0.741
Length, mm; median (range)	6 (1 - 20)	8 (2 -16)	0.920
AP diameter, mm; median (range)	9 (5 - 16)	9 (5 - 19)	0.905
Width, mm; mean ± SD	14 ± 5.0	15 ± 4.7	0.606
	n = 18	n = 28	
Bishop score; mean ± SD	5.8 ± 1.98	4.3 ± 1.76	0.006

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; AP, anterior-posterior; gws, gestational weeks.

Table 2. Gestational age, parity, ultrasound results, and Bishop score for women with spontaneous **onset** of labor ≤ 48 h vs. > 48 h after ultrasound examination

	Onset of labour		
	≤ 48 h	> 48 h	P-value
	n =35	n = 25	
Multipara; n (%)	20 (57%)	9 (36%)	0.104
< 42+0 gws at examination; n (%)	13 (37%)	6 (24%)	0.276
3D ultrasound			
Cervix			
Length, mm; median (range)	9 (2 - 4)	20 (5 - 39)	0.005
AP diameter, mm; mean ± SD	39 ± 8.3	35 ± 9.7	0.138
Width, mm; mean ± SD	44 ± 5.6	42 ± 5.8	0.335
Volume, cm^3; mean ± SD	27.4 ± 12.50	34.5 ± 13.19	0.037
Flow indices			
VI; median (range)	6.9 (0.3 - 23.7)	4.6 (0.9 - 16.4)	0.096
FI; mean ± SD	29.5 ± 3.36	29.0 ± 4.35	0.635
VFI; median (range)	1.8 (0.1 - 8.3)	1.0 (0.2 - 5.6)	0.116
Funnel; n (%)	20 (57%)	16 (64%)	0.220
Length, mm; median (range)	5 (1 - 20)	8 (2 - 16)	0.333
AP diameter, mm; median (range)	7 (5 - 18)	9 (5 - 19)	0.685
Width, mm; mean ± SD	15 ± 4.7	14 ± 4.9	0.716
	n = 25	n = 21	
Bishop score; mean ± SD	5.6 ± 1.96	4.1 ± 1.73	0.009

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; AP, anterior-posterior; gws, gestational weeks.

Table 3. Gestational age, parity, ultrasound results, and Bishop score for women who delivered vaginally ≤ 48 h vs. > 48 h after ultrasound examination

	Delivery ≤ 48 h	Delivery > 48 h	P-value
	n =29	n = 27	
Multipara; n (%)	17 (59%)	9 (33%)	0.056
< 42+0 gws at examination	9 (31%)	7 (26%)	0.672
3D ultrasound			
Cervix			
Length, mm; median (range)	8 (2 - 40)	19 (2 - 39)	0.015
AP diameter, mm; mean ± SD	39 (23 - 59)	37 (20 - 56)	0.330
Width, mm; mean ± SD	43 ± 5.9	43 ± 5.9	0.902
Volume, cm^3; mean ± SD	27.0 ± 12.74	33.2 ± 12.95	0.078
Flow indices			
VI; median (range)	6.9 (0.3 - 23.7)	4.3 (0.9 - 16.4)	0.118
FI; mean ± SD	28.9 ± 3.18	29.8 ± 4.43	0.334
VFI; median (range)	1.8 (0.1 - 8.3)	1.1 (0.2 - 5.6)	0.210
Funnel; n (%)	16 (55%)	18 (67%)	0.378
Length, mm; median (range)	6 (1 - 2)	7 (2 -16)	0.646
AP diameter, mm; median (range)	9 (5-18)	9 (5-19)	0.495
Width, mm; mean ± SD	15 ± 4.9	14 ± 4.4	0.314
	n = 22	n = 20	
Bishop score; mean ± SD	5.9 ± 1.79	3.8 ± 1.71	0.003

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; AP, anterior-posterior; gws, gestational weeks.

Table 4 Gestational age, parity, ultrasound results, and Bishop score for women who delivered vaginally ≤ 60 h vs. > 60 h after ultrasound examination

Results	Delivery ≤ 60 h	Delivery > 60 h	P-value
	n =34	n = 22	
Multipara; n (%)	19 (56%)	7 (32%)	0.075
< 42+0 gws at examination	12 (35%)	4 (19%)	0.158
3D ultrasound			
Cervix			
Length, mm; median (range)	9 (2 - 40)	19 (2 - 39)	0.015
AP diameter, mm; mean ± SD	39 (23 - 59)	34 (20 - 56)	0.220
Width, mm; mean ± SD	44 ± 6.0	42 ± 6.0	0.330
Volume, cm^3; mean ± SD	27.6 ± 12.65	33.7 ± 13.15	0.083
Flow indices			
VI; median (range)	7.1 (0.3 - 23.7)	3.5 (0.9 - 16.4)	0.050
FI; mean ± SD	29.3 ± 3.37	29.2 ± 4.52	0.877
VFI; median (range)	1.9 (0.1 - 8.3)	1.0 (0.2 - 5.6)	0.076
Funnel, n (%)	20 (59%)	14 (64%)	0.718
Length, mm; median (range)	5 (1 - 20)	8 (2 - 16)	0.466
AP diameter, mm; median (range)	7 (5 - 18)	9 (5 - 19)	0.900
Width, mm; mean ± SD	15 ± 4.7	14 ± 4.5	0.366
	n = 24	n = 18	
Bishop score; mean ± SD	5.6 ± 1.97	3.8 ± 1.74	0.004

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; AP, anterior-posterior; gws, gestational weeks.

Table 5. Logistic regression models to predict time to start of labor and delivery

	Maximum Likelihood Estimates		Odds Ratio Estimates		
	Standard Estimate	P-value	Effect	Point Estimate	95% Confidence Limits
Onset of labor > 24h					
Model building including Bishop score (n = 46)					
Intercept	-1.414	0.040			
Cervical length	0.149	0.008	Cervical length	1.160	1.040 – 1.295
Model building not including Bishop score (n = 60)					
Intercept	-7.409	0.012			
Cervical length	0.129	0.004	Cervical length	1.138	1.042 – 1.243
Cervical width	0.147	0.020	Cervical width	1.158	1.024 – 1.310
Onset of labor > 48h					
Model building including Bishop score (n = 46)					
Intercept	0.344	0.749			
Cervical length	0.107	0.008	Cervical length	1.113	1.028 – 1.204
Parity	-1.447	0.047	Parity	0.235	0.057 – 0.979
Model building not including Bishop score (n = 60)					
Intercept	-1.529	0.004			
Cervical length	0.076	0.009	Cervical length	1.079	1.020 – 1.142

Cont.

Table 5. Continued

	Maximum Likelihood Estimates		Odds Ratio Estimates		
	Standard Estimate	P-value	Effect	Point Estimate	95% Confidence Limits
Delivery > 48h					
Model building including Bishop score (n = 42)					
Intercept	3.424	0.006			
Bishop score	-0.722	0.003	Bishop score	0.486	0.303 – 0.780
Model building not including Bishop score (n = 56)					
Intercept	-6.019	0.036			
Cervical length	0.066	0.037	Cervical length	1.068	1.004 – 1.136
VI	-0.154	0.045	VI	0.857	0.737 – 0.996
FI	0.203	0.047	FI	1.225	1.003 – 1.495

Cont.

Table 5. Continued

	Maximum Likelihood Estimates		Odds Ratio Estimates		
	Standard Estimate	P-value	Effect	Point Estimate	95% Confidence Limits
Delivery > 60h					
Model building including Bishop score (n = 42)					
Intercept	0.621	0.608			
Bishop score	-0.442	0.033	Bishop score	0.643	0.428 – 0.966
Cervical length	0.082	0.048	Cervical length	1.086	1.001 – 1.178
Model building not including Bishop score (n =56)					
Intercept	-1.638	0.004			
Cervical length	0.076	0.011	Cervical length	1.079	1.017 – 1.144

VI, vascularity index; FI, flow index

Table 6. Areas under receiver operator characteristic (ROC) curves for Bishop score, cervical length and logistic regression models, and optimal cutoff values with regard to predicting start of labor >24 h and > 48h and delivery > 48h and >60h, and the sensitivity, false-positive rate (1 minus specificity), and positive and negative likelihood ratios for the optimal cutoff

	Area under ROC curve		Optimal cutoff	Sensitivity	False-positive rate	LR+	LR–
	Estimate	95% CI					
Start of labor > 24h							
For women with results for Bishop score (n = 46)[#]							
1st regression model[1]	0.839	0.721 – 0.958	0.56*	0.82	0.17	4.8	0.21
Cervical length	0.789	0.656 – 0.921	8.5 mm*	0.82	0.33	2.5	0.27
			12.5 mm*	0.64	0.11	5.8	0.40
Bishop score	0.724	0.569 – 0.880	5**	0.57	0.22	2.6	0.55
For all women (n =60)[#]							
1st regression model[1]	0.836	0.724 – 0.948	0.56*	0.85	0.19	4.5	0.19
Cervical length	0.758	0.628 – 0.889	8.5 mm*	0.85	0.38	2.2	0.25
			12.5 mm*	0.64	0.14	4.5	0.42

Cont.

Table 6. Continued

	Area under ROC curve		Optimal cutoff	Sensitivity	False-positive rate	LR+	LR–
	Estimate	95% CI					
Start of labor > 48h							
For women with Bishop score (n = 46)[#]							
2nd regression model [2]	0.788	0.659 – 0.917	0.44*	0.71	0.24	3.0	0.38
Cervical length, mm	0.783	0.646 – 0.920	8.5 mm*	0.91	0.40	2.3	0.16
			12.5 mm*	0.71	0.20	3.6	0.36
Bishop score	0.717	0.568 – 0.867	5**	0.76	0.36	2.1	0.37
For all women (n =60)[#]							
2nd regression model[2]	0.753	0.631 – 0.875	0.44	0.72	0.29	2.5	0.39
Cervical length	0.745	0.619 – 0.871	8.5 mm**	0.92	0.51	1.9	0.16

Cont.

Table 6. Continued

	Area under ROC curve		Optimal cutoff	Sensitivity	False-positive rate	LR+	LR–
	Estimate	95% CI					
Delivery > 48h							
For women with Bishop score (n = 42)[#]							
Bishop score, mm	0.816	0.682 – 0.950	5**	0.85	0.27	3.1	0.21
3rd regression model [3]	0.805	0.659 – 0.950	0.39*	0.85	0.23	3.7	0.19
Cervical length, mm	0.749	0.590 – 0.908	8.5 mm*	0.90	0.41	2.2	0.17
			12.5 mm*	0.70	0.18	3.8	0.37
For all women (n = 56)[#]							
3rd regression model [3]	0.784	0.654 – 0.914	0.56*	0.70	0.14	5.1	0.34
Cervical length	0.736	0.596 – 0.876	8.5 mm*	0.93	0.48	1.9	0.14
			12.5 mm*	0.74	0.21	3.6	0.33

Cont.

Table 6. Continued

	Area under ROC curve		Optimal cutoff	Sensitivity	False-positive rate	LR+	LR–
	Estimate	95% CI					
Delivery > 60h							
For women with Bishop score (n = 42)[#]							
4th regression model [4]	0.806	0.668 – 0.943	0.38	0.83	0.25	3.3	0.22
Cervical length	0.786	0.643-0.929	8.5 mm*	0.94	0.42	2.3	0.10
Bishop score	0.752	0.602 – 0.903	5**	0.83	0.33	2.5	0.25
For all women (n =56)							
Cervical length	0.748	0.618 – 0.878	8.5 mm*	0.96	0.53	1.8	0.09

ROC, receiver operating characteristics; CI, confidence interval; LR+ , positive likelihood ratio; LR–, negative likelihood ratio

The areas under the ROC curves to predict the outcome in these women do not differ significantly

*Larger values of the test result indicate stronger evidence for the outcome

** Smaller values of the test result indicate stronger evidence for the outcome

[1] Probability of start of labor >24 h = $[e^z/(1+e^z)]$ where z = -7.409 + 0.129 x cervical length + 0.147 x cervical width

Table 6. Continued

	Area under ROC curve	Optimal cutoff	Sensitivity	False-	LR+	LR–

	Estimate	95% CI	positive rate

[2] Probability of start of labor > 48 h = $[e^z/(1+e^z)]$ where z = 0.344 + 0.107 x cervical length -1.447 x parity

[3] Probability of delivery > 48 h = $[e^z/(1+e^z)]$ where z = -6.019 + 0.066 x cervical length - 0.154 x VI + 0.203 x FI

[4] Probability of delivery > 60 h = $[e^z/(1+e^z)]$ where z = 0.621 - 0.442 x Bishop score + 0.082 x cervical length