

OBSTETRICS

The short-term prediction of preterm birth: a systematic review and diagnostic metaanalysis

Amy B. Boots, DO; Luis Sanchez-Ramos, MD; Dawn M. Bowers, MD; Andrew M. Kaunitz, MD; Javier Zamora, PhD; Peter Schlattmann, MD, PhD

OBJECTIVE: To assess the diagnostic accuracy of fetal fibronectin (fFN), fetal breathing movements (FBM), and cervical length (CL) for the short-term prediction of preterm birth in symptomatic patients.

STUDY DESIGN: Diagnostic metaanalysis using bivariate methods.

RESULTS: Pooled sensitivities for fFN, FBM, and CL for delivery within 48 hours of testing were 0.62 (95% confidence interval [CI], 0.43–0.78), 0.75 (95% CI, 0.57–0.87) and 0.77 (95% CI, 0.54–0.90), respectively. Pooled specificities for fFN, FBM, and CL for delivery within 48 hours were 0.81 (95% CI, 0.74–0.86), 0.93 (95% CI, 0.75–0.98) and 0.88 (95% CI, 0.84–0.91). Pooled sensitivities for fFN, FBM, and CL for delivery within 7 days were 0.75 (95% CI, 0.69–0.80), 0.67 (95% CI, 0.43–0.84), and 0.74 (95% CI, 0.58–0.85). Pooled specificities for fFN, FBM, and CL for delivery

within 7 days were 0.79 (95% CI, 0.76–0.83), 0.98 (95% CI, 0.83–1.00) and 0.89 (95% CI, 0.85–0.92). Based on a pretest probability of 10% for delivery within 48 hours, posttest probabilities (positive and negative) were 27% and 5% for fFN, 54% and 3% for FBM, and 42% and 3% for CL. For a pretest probability of 20% for delivery within 7 days, posttest probabilities (positive and negative) were 48% and 7% for fFN, 89% and 8% for FBM, and 63% and 7% for CL.

CONCLUSION: In symptomatic patients, for fFN, absence of FBM, and CL have diagnostic use as predictors of delivery within 48 hours and within 7 days of testing. Absence of FBM appears to be the best test for predicting preterm birth.

Key words: cervical length, fetal breathing movements, fetal fibronectin, preterm birth

Cite this article as: Boots AB, Sanchez-Ramos L, Bowers DM, et al. The short-term prediction of preterm birth: a systematic review and diagnostic metaanalysis. *Am J Obstet Gynecol* 2014;210:54.e1-10.

Preterm birth is the leading cause of neonatal morbidity and mortality in developed countries.¹⁻³ Although there have been small decreases during the past 4 years, the incidence of preterm birth in the United States remains at 12%, higher than in many developed countries.⁴

Preterm labor has emerged as one of the most common and costly obstetric indications for hospital admissions and unscheduled visits; a trend that may reflect awareness of our high rate of preterm birth.⁵ In a population-based US report, suspected preterm labor accounted for nearly 33% of the admissions to an inpatient prenatal unit.⁶ However, the

diagnosis of preterm labor is subjective and unreliable and varies substantially among published studies.⁷ In a meta-analysis assessing the diagnostic accuracy of cervicovaginal fetal fibronectin (fFN), the prevalence of preterm delivery within 7 days of testing ranged from 1.8% to 29.7% with a median prevalence of 8%.⁸ A New York City hospital reported that 80% of women admitted for preterm labor required no treatment.⁹ In a multicenter US trial, only 3% of women with suspected preterm labor who were included in studies of fFN testing delivered within 7 days.¹⁰ Randomized clinical trials involving tocolytic therapy found

that nearly 50% of the study population delivered at term.¹¹⁻¹³ Hospital admissions for patients who are not in true preterm labor often result in the use of unnecessary and potentially harmful medications including tocolytics, corticosteroids, and antibiotics. In addition, hospitalization is expensive, an important source of stress, disruptive for the family, and often has a negative psychologic impact.^{14,15}

Cervicovaginal fFN, fetal breathing movements (FBM), and transvaginal sonographic cervical length (CL) measurements (TVS) represent readily available and well-tolerated testing modalities. Although several systematic reviews have assessed fFN, FBM, and TVS for the short-term prediction of preterm birth in women with signs and symptoms of preterm labor,¹⁶⁻¹⁹ we are not aware of a comprehensive systematic review comparing the accuracy of the 3 techniques. Although a recently published systematic review collated information about numerous tests potentially useful in the short-term prediction of preterm birth in women with signs and symptoms of

From the Department of Obstetrics and Gynecology (Drs Boots, Sanchez-Ramos, Bowers, and Kaunitz), University of Florida College of Medicine, Jacksonville, FL; the Clinical Biostatistics Unit, Ramon y Cajal Hospital, Madrid, and CIBER Epidemiología y Salud Pública (CIBERESP) (Dr Zamora), Barcelona, Spain; and the Institute of Medical Statistics, Computer Sciences and Documentation (Dr Schlattmann), Jena University Hospital, Friedrich Schiller University, Jena, Germany.

Received June 13, 2013; revised Aug. 8, 2013; accepted Sept. 6, 2013.

The authors report no conflict of interest or financial support.

Reprints: Amy B. Boots, DO, Department of Obstetrics and Gynecology, University of Florida, College of Medicine-Jacksonville, 653-1 West 8th St., Jacksonville, FL 32209. amy.boots@jax.ufl.edu.

0002-9378/\$36.00 • © 2014 Mosby, Inc. All rights reserved. • <http://dx.doi.org/10.1016/j.ajog.2013.09.004>

preterm labor, this publication did not assess the diagnostic accuracy of these tests in depth.²⁰ There is a need to rationalize the use of testing to predict preterm birth in women with signs and symptoms of preterm labor, and clear summaries of evidence can assist in the generation of practical guidance.

The objective of this systematic review is to provide a comprehensive and up-to-date overview of diagnostic accuracy of cervicovaginal fFN, absence of FBMs, and TVS CL measurement for the short-term prediction of delivery in patients with signs and symptoms of preterm labor. The selected outcomes of delivery within 48 hours and/or within 7 days were chosen as the reference standard.

MATERIALS AND METHODS

The proposed systematic review with metaanalysis was preceded by a detailed study protocol stating the question to be addressed, the subgroups of interest, the methods, and criteria to be used for identifying and selecting relevant studies and extracting and analyzing information. This systematic review and metaanalysis was conducted according to the Metaanalysis of Observational Studies in Epidemiology (MOOSE) guidelines.²¹ We also followed guidelines for meta-analyses and systematic reviews evaluating screening and diagnostic tests²² and adopted similar standards for the reporting of accuracy studies (STARD).²³

Using available computerized databases, references of published systematic reviews, syllabi from scientific meetings, and chapters from textbooks, we identified studies that evaluated the diagnostic accuracy of cervicovaginal fFN, absence of FBM, and TVS measurement of CL for the short-term prediction of preterm birth in symptomatic women. We included only retrospective or prospective cohort studies that reported rates of preterm delivery within 48 hours and within 7 days of testing. The computerized databases searched were: MEDLINE (national Library of Medicine, Bethesda, MD), EMBASE (Elsevier Science, New York, NY), Current Contents (Institute for Scientific Information, Philadelphia, PA), Silver Platter (Silver Platter Information Inc, Norwood MA), and the

Cochrane Library (Cochrane). The searches were conducted for literature published between 1966 and April 2013. Non-English language articles were included and translated before data abstraction. Abstracts from relevant scientific meetings that took place since 1981 were hand searched for unpublished studies or abstracts. We cross-referenced articles obtained from the database search and contacted experts in the field to uncover additional unpublished articles or abstracts. If a particular patient population was reported in more than 1 publication, we selected the article that provided the most complete data set. Medical subject heading search terms included “diagnostic accuracy,” “preterm delivery,” “sensitivity and specificity,” “likelihood ratios,” “sonography,” “ultrasound,” “fetal fibronectin,” “fetal breathing movements,” and “cervical length.”

Studies were selected for review if they included symptomatic pregnant women who after spontaneous onset of labor underwent either cervicovaginal fFN testing, assessment of FBM, or TVS measurement of CL before 37 weeks’ gestation. Studies evaluating the performance of cervicovaginal fFN were selected if they included known gestational ages after spontaneous labor and delivery and used preterm delivery within 48 hours and/or within 7 days as the reference standard. Studies describing the presence or absence of FBM were selected if they examined the accuracy of the absence of FBM in short-term prediction of spontaneous preterm birth. The reference standard was delivery within 48 hours and/or within 7 days in women with threatened preterm labor, intact fetal membranes, without antepartum bleeding, and known gestation at spontaneous birth. FBM, in general, were defined as present if sustained for at least 20 seconds and absent if no sustained movements were detected over an observation period of at least 30 minutes. Studies evaluating the diagnostic accuracy of TVS were selected if they assessed the short-term prediction of preterm birth. The reference standard was delivery within 48 hours and/or within 7 days in women with intact membranes and the spontaneous onset of labor for a

single or various cutoff values of CL. Studies that included data from singleton and multifetal pregnancies were analyzed separately and double counting was avoided whenever possible. Studies included for review were selected by 2 of the authors (A.B. and L.S.R.) after scrutinizing the electronic searches and obtaining manuscripts of all citations that met the inclusion criteria cited above. Discrepancies were discussed with a third author (A.M.K.) and resolved by consensus. The main outcomes assessed and the data abstraction forms were finalized before analysis.

Study characteristics, quality of study design, and accuracy of results were obtained from each of the selected studies. To assess methodologic quality, we adapted criteria from a tool developed for the quality assessment of studies of diagnostic accuracy included in systematic reviews (QUADAS).²⁴ The revised criteria include 14 items covering several dimensions of study quality: patient spectrum, reference standard, verification bias, review bias, clinical review bias, incorporation bias, masking, test execution, study withdrawals, and indeterminate results. Studies were of high quality when at least 12 items were met; moderate when at least 10 items were met and low for the remainder (<10).

We abstracted data for women with signs and symptoms of preterm labor and spontaneous preterm delivery within 48 hours and/or 7 days of testing with each modality. Statistical analyses were performed according to current recommendations.^{25,26} Accuracy data were used to construct 2 × 2 contingency tables of fFN, FBM, and TVS CL results and the diagnosis of preterm delivery within 48 hours and within 7 days of testing. The true positive, false positive, true negative, and false negative values were abstracted and recorded. True positives included those that delivered within 48 hours and/or within 7 days with a positive fFN, absent FBM, or a short CL. In some cases, these data were not provided in the original publications and had to be calculated from the raw data or data obtained after contacting the authors.

Study heterogeneity was assessed in several ways. We used paired forest plots

of sensitivity and specificity to represent individual studies' estimates along with their precision, represented by their exact 95% confidence intervals (CIs). As a potential cause of heterogeneity in sensitivity and specificity among the included studies, threshold/cut off effect was tested with the Spearman correlation coefficient between the logit of sensitivity and logit of 1-specificity. Heterogeneity induced by factors other than threshold/cutoff effect was evaluated by using Cochran-Q and the inconsistency index statistic (I^2).^{27,28} For the latter, value of 0% indicates no observed heterogeneity, and values greater than 50% may be considered substantial heterogeneity. Statistical significance of heterogeneity testing was assumed when a *P* value was less than .10.

Studies that reported the results using different fFN assays or various cutoff values for CL were analyzed initially as a single group. We also conducted separate metaanalyses within predetermined subgroups to assess whether diagnostic accuracy varied according to type of fFN assay and the CL cutoff values used.

Pooled measures of sensitivity and specificity were determined after first estimating the underlying summary receiver operating characteristic (SROC) curve using a bivariate random effects model and then producing averaged estimations of diagnostic accuracy indices along with their corresponding 95% confidence ellipses. This model acknowledges the difference in precision by which sensitivity and specificity have been measured in each study owing to the differences in the number of women with or without preterm deliveries. We aggregated accuracy data of TVS using the cutoff selected by each individual author even though these cutoffs varied. Because we noted a threshold effect when reviewing accuracy estimates, we then fitted an overall SROC curve. This SROC curve was summarized by the area under the curve along with its 95% CI. To provide meaningful sensitivity and specificity estimates for TVS CL, we pooled data from those studies that used a 15 mm cutoff point. Summary estimates of accuracy indices were then obtained. The diagnostic accuracy of the 3 tests was compared by introducing a dummy variable coding for

the test modality into the bivariate model. This approach allowed us to check for differences in sensitivity and specificity among the 3 diagnostic modalities.

To analyze the effect of different study characteristics on diagnostic accuracy, we fit several univariate metaregression models as described by Lijmer et al.²⁹ The diagnostic measure used for metaregression was the relative diagnostic odds ratio (RDOR).³⁰ The variables assessed by metaregression for the fFN studies included: method of testing (enzyme-linked immunosorbent assay [ELISA] compared with others), blinding of tests results, quality of studies (scores of at least 10 points), prevalence of delivery within 7 days, country of origin (United States compared with others), year of publication (older than 2002 compared with more recent), and sample size. Variables assessed for metaregression of the TVS CL studies included: CL cutoff, blinding of study, quality of studies (scores of at least 10 points), prevalence of delivery within 7 days, prospective study and sample size. Owing to the small number of studies, metaregression was not performed for the FBM studies nor was multivariate modeling performed for any of the testing modalities.

Posttest probabilities were calculated based on 1 common pretest probability (10%) of delivery within 48 hours and (20%) for delivery within 7 days. Following current guidelines, we evaluated for small study effects potentially because of publication bias using Deek's funnel plot asymmetry test ($P < .10$ considered significant),³¹ and by visually inspecting produced funnel plots. All analyses were performed using Stata 11.0 statistical software (StataCorp, College Station, TX) and MetaDisc 1.4 (www.hrc.es/investigacion/metadisc_en.htm).

RESULTS

The process of study selection revealed a total of 1371 studies (Figure 1). In screening titles and abstracts we concluded that 1075 were irrelevant, leaving 296 studies for potential inclusion. After reviewing the full text, 224 studies were excluded, leaving 72 studies for inclusion (38 fFN,³²⁻⁶⁹ 10 FBM,⁷⁰⁻⁷⁸ and 24

TVS^{51,52,54,60,62,63,65,66,69,79-93}). Several reports assessed both the diagnostic accuracy of fFN and sonographic measurement of CL.^{51,52,54,60,62,63,65,66,69}

The 38 fFN studies were published between 1995 and 2013, and the number of patients analyzed ranged from 25 to 725 (6383 women in aggregate). These publications originated from the United States (15), France (4), United Kingdom (2), Germany (2), Japan (2), Netherlands (2), Turkey (2) with single publications from Chile, Canada, South Africa, New Zealand, Italy, Ecuador, Singapore, Mexico, and Australia.

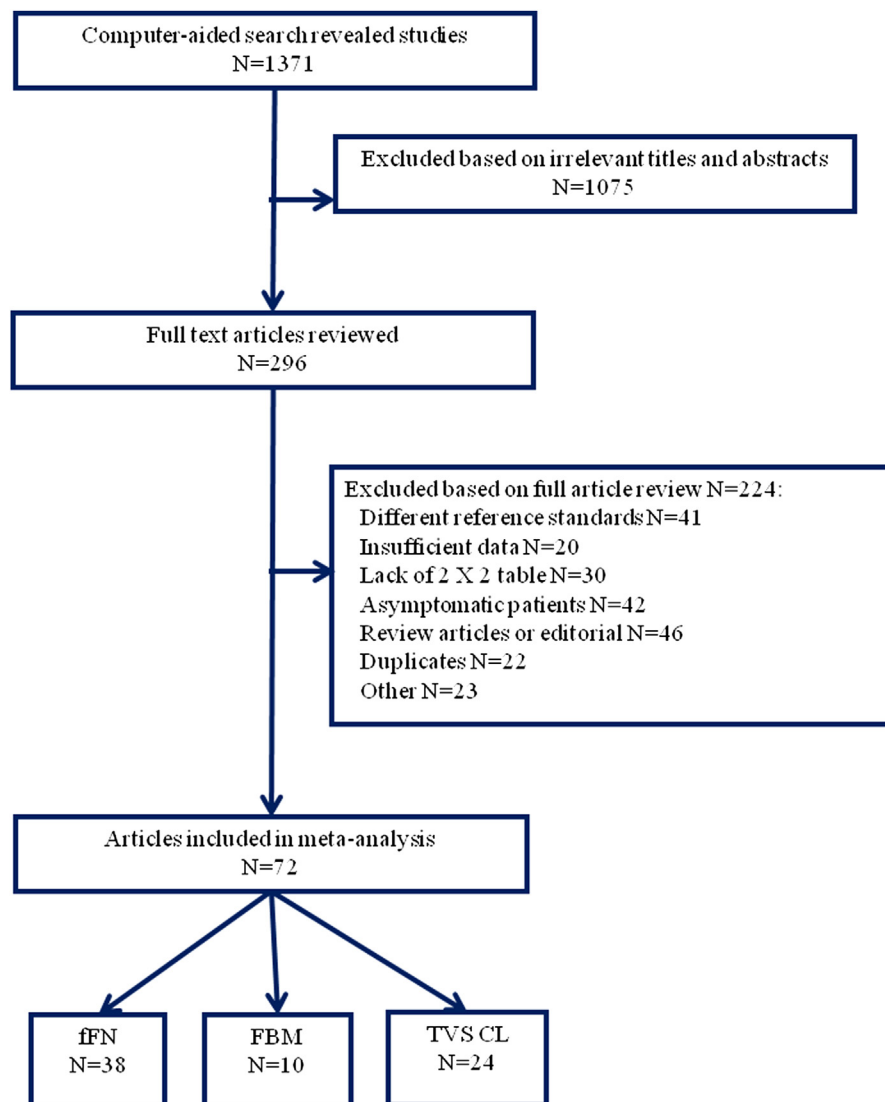
The 10 FBM studies were published between 1983 and 2001, and the number of patients analyzed ranged from 24 to 70 (391 women in aggregate). These publications originated from the United States (3), United Kingdom (3), Israel (2), with single studies from Ireland, and Poland.

The 24 TVS CL studies were published between 2005 and 2013, and the sample sizes ranged from 29 to 559 (5112 women in aggregate). The TVS originated from Italy (3), France (3), Chile (2), Greece (2), Denmark (2), Turkey (2), Netherlands (2), Spain (2) with single studies from United Kingdom, South Africa, Tunisia, Japan, India, and Mexico. Overall, authors of 9 reports were contacted and asked for additional data or clarification of their results.^{52,67,69,83-86,88,93} The QUADAS scores varied from 9 to 14 (median 12) in fFN studies, 11 to 12 (median 11) in FBM studies, and 8 to 12 (median 10) in TVS CL studies.

Heterogeneity was assessed for each of the 3 testing modalities (delivery within 48 hours and/or within 7 days). Of the 3 testing modalities, TVS, using the selected cutoff used by each author, was the only test with evidence of a significant threshold effect (correlation coefficient 0.481; $P = .02$) based on Spearman's correlation coefficient.

Table 1 depicts the diagnostic accuracy of fFN, FBM and TVS CL measurements in predicting spontaneous delivery within 48 hours and within 7 days of testing. For the 48 hour analysis, data were available for 4 fFN, 4 FBM studies, and 9 TVS CL studies. For both spontaneous, delivery within 48 hours and within 7 days, metaanalysis was based on

FIGURE 1
Flowchart of studies included in the metaanalysis



Boots. Short-term prediction of preterm birth. *Am J Obstet Gynecol* 2014.

the bivariate random-effects model in the presence of significant heterogeneity. For the prediction of delivery within 48 hours, the pooled sensitivity was higher for TVS; however, the pooled specificity was higher for FBM. Both the likelihood ratio for a positive test and the diagnostic odds ratio were higher for FBM than for TVS. However, the likelihood ratios for a negative test were similar for the 2 tests. For the prediction of delivery within 7 days, data were available for 38 fFN, 7 FBM studies, and 24 TVS CL studies. The pooled sensitivity for fFN was highest and pooled specificity was again

higher for TVS than for FBM; however, pooled specificity was higher for FBM. Both the likelihood ratio for a positive test and the diagnostic odds ratio were higher for FBM than for fFN and TVS. However, the likelihood ratio for a negative test was lowest for TVS and similar for fFN and FBM.

With respect to TVS, measures of diagnostic accuracy were also performed using several CL cutoffs to determine which of them best predicted delivery within 7 days of testing (Table 2). Using a cutoff measurement of ≤ 15 mm appeared most accurate in predicting

spontaneous delivery within 7 days and was associated with a sensitivity and specificity of 0.74 (95% CI, 0.58–0.85) and 0.89 (95% CI, 0.85–0.92), respectively. Increasing CL cutoff was positively correlated with sensitivity and negatively correlated with specificity.

The comparisons of sensitivity and specificity amongst the 3 tests showed no statistically significant differences in predicting delivery within 48 hours and within 7 days. However, FBM at 48 hours showed statistically significant better specificity than fFN and TVS ($P = .024$ and $P = .046$ for fFN and TVS, respectively). Similarly, the specificity of FBM at 7 days was statistically greater than the other 2 modalities ($P = .0000$ and $P = .0005$ for fFN and TVS, respectively).

We plotted hierarchical SROC curves to graphically present the results of all 3 testing modalities. Figure 2, A-D, representing fFN, FBM, and TVS (select cutoffs and ≤ 15 mm), shows the ROC place with the fitted SROC curve and summary accuracy estimates along with the corresponding 95% confidence prediction ellipses. In hierarchical SROC curves, the index test's sensitivity (true positive rate) was plotted on the y axis against 1-specificity (false negative rate) on the x axis. In addition, the 95% confidence region and a 95% prediction region around the pooled estimates were plotted to illustrate the precision with which the pooled values were estimated (confidence ellipse of a mean) and to show the amount of between study variation (prediction ellipse; the likely range of values for a new study).⁹⁴ The area under the curve (AUC) and 95% CI for delivery within 48 hours and within 7 days were also calculated and the results are shown in Table 1. The results for FBM was 0.91 (95% CI, 0.88–0.93) for delivery within 7 days. There were no differences in the overall accuracy of the 3 tests.

Pre- and posttest probabilities for the prediction of delivery within 48 hours and/or within 7 days were reported for cervicovaginal fFN, absence of FBM and TVS CL measurement using 1 common pretest probability of delivery within 48 hours (10%) and within 7 days (20%), and are represented in Table 3. Because of low negative LRs, presence of FBM

TABLE 1
Summary estimates for the short-term prediction of delivery among the tests

Variable	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	DOR (95% CI)	AUC (95% CI)
Delivery within 48 hours						
fFN	0.62 (0.43–0.78)	0.81 (0.74–0.86)	3.3 (2.1–5.0)	0.47 (0.29–0.76)	7 (3–17)	0.74 (0.63–0.83)
FBM	0.75 (0.57–0.87)	0.93 (0.75–0.98)	10.4 (2.8–38.4)	0.27 (0.15–0.49)	37.8 (9–164)	0.83 (0.72–0.90)
TVS CL ^a	0.77 (0.54–0.90)	0.88 (0.84–0.91)	6.4 (4.7–8.7)	0.26 (0.12–0.58)	24 (9–65)	0.90 (0.88–0.93)
Delivery within 7 days						
fFN	0.75 (0.69–0.80)	0.79 (0.76–0.83)	3.6 (3.1–4.3)	0.31 (0.25–0.39)	11.5 (8–16)	0.84 (0.80–0.87)
FBM	0.67 (0.43–0.84)	0.98 (0.83–1.00)	31.6 (4.1–244)	0.34 (0.18–0.64)	93 (15–592)	0.91 (0.88–0.93)
TVS CL ^a	0.74 (0.58–0.85)	0.89 (0.85–0.92)	6.8 (5.1–9.2)	0.29 (0.17–0.49)	23 (12–46)	0.91 (0.67–0.98)

CI, confidence interval; CL, cervical length; DOR, diagnostic odds ratio; FBM, fetal breathing movements; fFN, fetal fibronectin; LR, likelihood ratio; TVS, transvaginal sonographic.

^a 15 mm cutoff.

Boots. Short-term prediction of preterm birth. *Am J Obstet Gynecol* 2014.

and normal CL are associated with substantial decreases in posttest probabilities. For example, the presence of FBM is associated with a reduction in the probability of delivery within 48 hours from 10% to a posttest probability of 3%. Similarly, the absence of FBM is associated with an increase in the probability from 20% to a posttest probability of 89% for delivery within 7 days.

Metaregression analysis for fFN found that test accuracy was affected by prevalence of delivery within 7 days of

testing, score, year of publication, and by blinding, with blinded studies observing greater test accuracy than unblinded studies. The analysis also found that studies published after 2002 were less accurate than studies published before 2002 (relative DOR, 0.37; 95% CI, 0.19–0.69). Metaregression analysis for the TVS CL studies found that test accuracy was affected by the type of study, with prospective studies finding greater test accuracy than retrospective studies (RDOR, 6.28; 95% CI, 1.65–23.92). Visual inspection of produced funnel

plots for the three testing modalities suggested evidence of small study effects for fFN potentially resulting from publication bias, an observation also supported by the results of Deek's funnel plot asymmetry test ($P = .05$).

COMMENT

Overall, our systematic review and meta-analysis found that although the number of studies and patients are limited, absence of FBM appears to afford the highest diagnostic accuracy in predicting preterm birth in symptomatic women. Among the

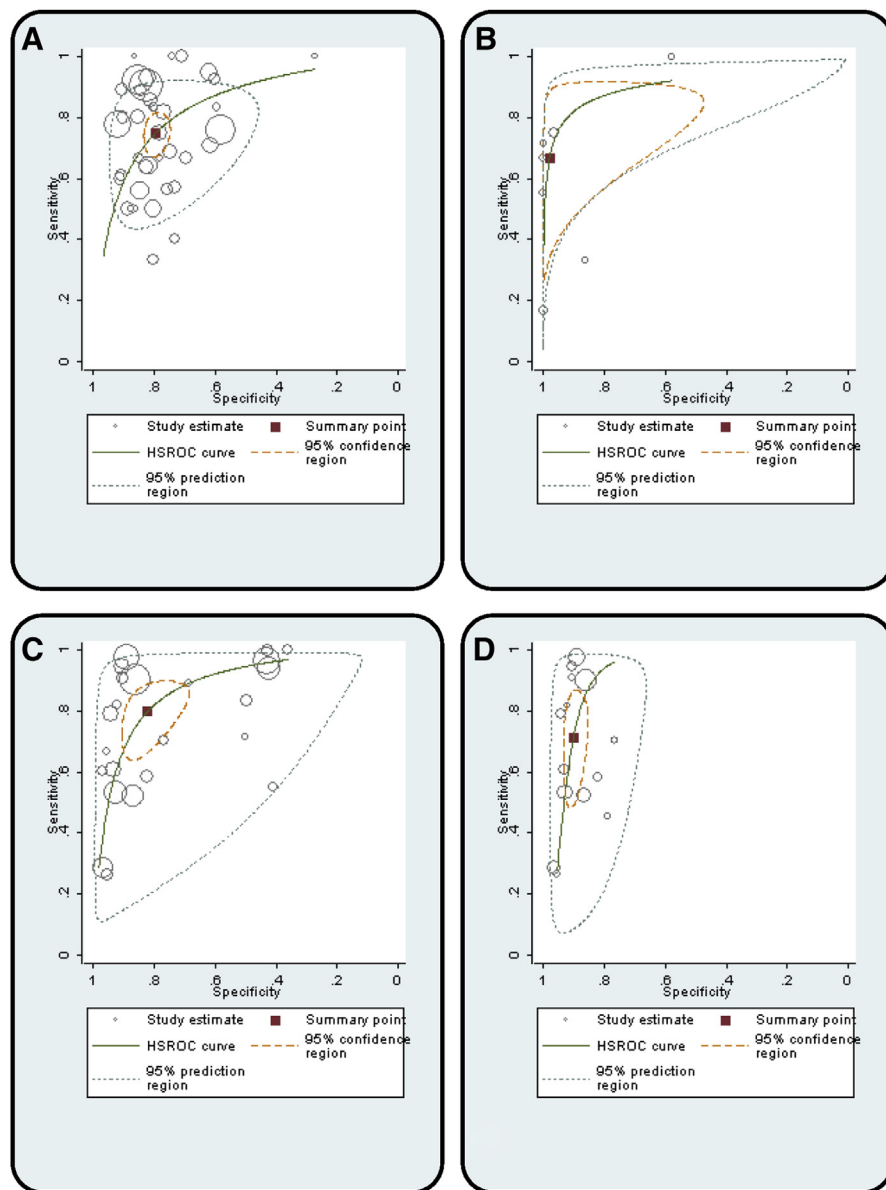
TABLE 2
Diagnostic accuracy of delivery within 7 days based on various TVS CL cutoffs

Cutoff	Studies, n	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	DOR (95% CI)
5 mm	2	0.35 (0.23–0.53)	0.99 (0.98–1.00)	21 (4–118)	0.72 (0.54–0.97)	29 (4–217)
10 mm	1	0.65 (0.49–0.79)	0.96 (0.94–0.97)	15.2 (9.4–24.6)	0.36 (0.24–0.55)	42 (19–90)
15 mm	15	0.74 (0.58–0.85)	0.89 (0.85–0.92)	6.8 (5.1–9.2)	0.29 (0.17–0.49)	23.4 (11.9–45.9)
20 mm	3	0.84 (0.48–0.97)	0.82 (0.60–0.93)	4.5 (2.2–9.4)	0.20 (0.06–0.73)	22 (7–76)
25 mm	9	0.78 (0.66–0.86)	0.70 (0.56–0.81)	2.6 (1.7–3.9)	0.32 (0.20–0.50)	8.1 (3.8–17)
30 mm	5	0.95 (0.90–0.97)	0.46 (0.42–0.50)	1.8 (1.6–1.9)	0.11 (0.06–0.21)	15.8 (8.3–30.1)
35 mm	1	1.00 (0.72–1.00)	0.36 (0.26–0.48)	1.5 (1.2–1.9)	0.11 (0.01–1.8)	13 (1–232)
All	37	0.79 (0.70–0.86)	0.82 (0.75–0.87)	4.4 (3.2–6.0)	0.25 (0.18–0.35)	17 (11.6–26)
Selected	24	0.81 (0.70–0.89)	0.82 (0.72–0.88)	4.4 (3.0–6.5)	0.23 (0.15–0.37)	19 (11–33)

All, all reported cutoffs; CI, confidence interval; CL, cervical length; DOR, diagnostic odds ratio; LR, likelihood ratio; Selected, cutoffs selected by the authors of each study; TVS, transvaginal sonographic.

Boots. Short-term prediction of preterm birth. *Am J Obstet Gynecol* 2014.

FIGURE 2
HSROC curves for the testing modalities



A, HSROC curve of fFN delivery within 7 days. AUC, 0.84 and 95% CI, 0.80–0.87. **B**, HSROC curve of FBM delivery within 7 days. AUC, 0.91 and 95% CI, 0.88–0.93. **C**, HSROC curve of TVS (select cutoff by each author) delivery within 7 days. AUC, 0.88 and 95% CI, 0.70–0.96. **D**, HSROC curve of TVS (15 mm cutoff) delivery within 7 days. AUC, 0.91 and 95% CI, 0.67–0.98.

AUC, area under the curve; CI, confidence interval; FBM, fetal breathing movements; fFN, fetal fibronectin; HSROC, hierarchical summary receiver operator characteristic; TVS, transvaginal sonographic cervical length measurements.

Boots. Short-term prediction of preterm birth. *Am J Obstet Gynecol* 2014.

3 tests, FBM was associated with specificity superior to that of fFN and TVS in predicting delivery within 48 hours of testing and delivery within 7 days although sensitivity was similar for all 3 tests.

A number of studies have compared the diagnostic accuracy of cervicovaginal

fFN and sonographic measurements of CL with the primary outcomes of preterm delivery before 37 weeks' gestation.^{95–98} Randomized clinical trials have found that tocolytic therapy results in pregnancy prolongation of 48 hours to 7 days, providing time for beneficial

measures, such as transfer to a level III perinatal center and administration of corticosteroids for the enhancement of lung maturation.^{99,100} Accordingly, because the morbidity of neonates born after 34 weeks' gestation has diminished, clinical management of women presenting with symptoms or signs of preterm labor involves estimating the risk of imminent preterm delivery. Therefore, we selected delivery within 48 hours and/or within 7 days as the primary outcome measures as they appear to be more clinically relevant than preterm delivery greater than 7 days after testing.

We reviewed the literature to appraise and compare the diagnostic accuracy of fFN, FBM, and TVS for the short-term prediction of spontaneous preterm birth in women with signs and symptoms of preterm labor. Bivariate meta-analyses of the available data found a robust association between absent FBM with delivery within 7 days of testing. In contrast, fFN and TVS were noted to have limited and moderate accuracy, respectively, for the short-term prediction of delivery in women with signs and symptoms of preterm labor.

The clinical usefulness of a diagnostic test is largely determined by the accuracy with which it identifies its target disorder, and this accuracy is best determined by assessing likelihood ratios. As a general guide, results of tests with positive likelihood ratios greater than 10 or negative likelihood ratios less than 0.1 are associated with large and often conclusive shifts from pretest to posttest probability.¹⁰¹ The positive LRs for predicting delivery within 48 hours and/or within 7 days of testing for FBM were 10.4 and 31.6, respectively. In addition, based on the actual pretest probability of delivery for each test, negative likelihood ratios associated with FBM for predicting delivery at greater than 48 hours and greater than 7 days were 0.27 and 0.34 respectively. To enhance clinical applicability, we estimated posttest probabilities based on one common pretest probability of 10% for delivery within 48 hours and 20% for delivery within 7 days.

The findings of this systematic review indicate that the assessment of FBM in symptomatic patients permits the prediction of preterm delivery within 7 days

TABLE 3
Posttest-test probabilities based on common pretest probabilities

Variable	LR+ (95% CI)	LR- (95% CI)	Probability before test	Probability after positive test	Probability after negative test
Delivery within 48 hours: pretest probability (10%)					
fFN	3.3 (2.1–5.0)	0.47 (0.29–0.76)	10%	27%	5%
FBM	10.4 (2.8–38.4)	0.27 (0.15–0.49)	10%	54%	3%
TVS CL ^a	6.4 (4.7–8.7)	0.26 (0.12–0.58)	10%	42%	3%
Delivery within 7 days: pretest probability (20%)					
fFN	3.6 (3.1–4.3)	0.31 (0.25–0.39)	20%	48%	7%
FBM	31.6 (4.1–244)	0.34 (0.18–0.64)	20%	89%	8%
TVS CL ^a	6.8 (5.1–9.2)	0.29 (0.17–0.49)	20%	63%	7%

CI, confidence interval; CL, cervical length; FBM, fetal breathing movements; fFN, fetal fibronectin; LR, likelihood ratio; TVS, transvaginal sonographic.

^a 15 mm cutoff.

Boots. Short-term prediction of preterm birth. *Am J Obstet Gynecol* 2014.

with high accuracy. As an example, in the assessment of women with an overall 20% pretest probability of delivery within 7 days, the absence or presence of FBM is predictive of 89% and 8% likelihood of this outcome, respectively.

In a previous metaanalysis assessing the diagnostic accuracy of fFN for the short-term prediction of birth in symptomatic patients, we concluded that this test has limited accuracy.⁸ Pooled positive and negative LRs in that metaanalysis were 4.20 and 0.29, respectively. The findings of our current updated and revised metaanalysis with positive and negative LRs of 3.6 and 0.31, respectively, confirm that fFN assay is of limited accuracy in the short-term prediction of spontaneous delivery in symptomatic patients. A recent study evaluated a quantitative fFN test for spontaneous preterm birth in symptomatic women.¹⁰² The authors reported that quantitative information has added value over the currently used qualitative fFN test. Perhaps by increasing the fFN threshold, its value in predicting preterm birth in symptomatic patients will be enhanced.

In a 2004 metaanalysis, Honest et al¹⁶ indeed noted that the absence of FBM has the potential to be a useful test in predicting preterm birth within 48 hours as well as within 7 days of testing. However, this report included patients

with premature rupture of fetal membranes and antepartum hemorrhage. In addition, Honest et al¹⁶ limited quantitative assessment of diagnostic accuracy to simply pooling likelihood ratios and did not perform a bivariate metaanalysis as currently recommended for diagnostic studies.²⁶ Finally, this metaanalysis failed to include 2 studies that appear to have met their inclusion criteria.^{71,72} Each of these factors likely contributed to the modest differences in diagnostic accuracy for FBM noted in the Honest et al¹⁶ and the present metaanalyses.

Although the results of our metaanalysis indicate that assessment of FBM appears to represent a useful test to confirm and to rule out preterm labor, a note of caution is in order. Despite our inclusion of additional studies, we agree with Honest et al's¹⁶ critique regarding the small number of studies and women as well as the concerns expressed regarding quality of the methodology of the included studies.

Transvaginal sonographic cervical measurement as a test for the prediction of preterm birth in symptomatic women has been the subject of several studies and systematic reviews.^{17,19,98} Prior metaanalyses have focused on the predictive accuracy for delivery <34 and <35 weeks in symptomatic and asymptomatic women, and have been hampered by the small number

of studies included and several methodologic flaws. In our current systematic review, we included 24 studies for bivariate metaanalysis and noted that TVS appears to be accurate in excluding preterm labor, with negative likelihood ratios of 0.26 and 0.29 for predicting delivery within 48 hours and within 7 days, respectively. These ratios were lower than those observed for fFN and FBM. In assessing the diagnostic performance of TVS assessment of CL using various cutoffs, as expected, higher cutoffs were associated with greater sensitivity but decreasing specificity in predicting delivery within 7 days. A randomized controlled trial found that women with threatened preterm labor and CL more than 15 mm on TVS were able to safely avoid tocolytic and corticosteroid therapy as well as unnecessary hospital admission.¹⁰³

Several studies and a recent systematic review¹⁰⁴ have assessed the accuracy of combining fFN with TVS with the aim of improving the accuracy in predicting preterm delivery in symptomatic patients. Unfortunately, the number of studies with data amenable for comparison is small and several of the studies selected for comparison in the systematic review did not present the results from women who had both tests. A further limitation in attempting to draw clinical inferences regarding the combined use of fFN and TVS is that most of the

outcomes reported were for deliveries greater than 7 days after testing.

No published studies have assessed whether combining fFN and/or TVS with FBM might improve the short-term prediction of preterm birth in symptomatic women. Because of the low negative LR noted in the group of patients evaluated with TVS and the high positive LR of those assessed with FBM, combining these 2 readily available assessments might improve overall diagnostic accuracy.

This systematic review with diagnostic metaanalysis has strengths as well as limitations. We performed an exhaustive literature search without language restrictions and systematically assessed the methodologic quality of the studies. We included all available studies reporting on the accuracy of fFN, absent FBM, and sonographic measurement of CL for a variety of cutoff points in predicting preterm birth within 48 hours and within 7 days of testing. We did not exclude unpublished studies or those in languages other than English. We also contacted authors, experts, and the manufacturer of one of the reviewed index test. Study selection was conducted independently by 2 reviewers, as were data extraction and quality review. Other strengths of our review include our focus on women with signs and symptoms of preterm labor, large sample size, and performance of metaanalysis.

Unfortunately, in most of the studies included in this review the results of the tests used were not blinded to the managing obstetrician, and interventions (hospitalization, tocolysis, and corticosteroids) were often performed based on these results. However, because the use of tocolysis has only been shown to be effective at short-term prolongation of pregnancy, the strong predictive value of absent FBM and TVS CL measurement may allow the identification of a subgroup of women that can avoid hospital admissions and unnecessary and potentially harmful treatments.

As observed in the present metaanalysis, there is often a high degree of heterogeneity between the results of diagnostic accuracy studies. Heterogeneity may be because of differences in study populations, methodology, variations in diagnostic threshold, and quality of the studies.

One of the most important goals in conducting a metaanalysis is to explore the sources of heterogeneity. Exploration for sources of heterogeneity was performed by using metaanalysis. However, metaanalysis cannot always account for heterogeneity. This limitation can be overcome only by performing metaanalyses of individual participant data.¹⁰⁵

In conclusion, assessment of cervicovaginal fFN assay, FBM, and sonographic measurement of CL have use in identifying which women who present with signs and symptoms of preterm labor will benefit from admission to the hospital for corticosteroid, tocolytic therapy, and/or transfer to a facility which can provide a more advanced level of perinatal care. Similarly, these tests can detect women with signs and symptoms of preterm labor who can safely avoid hospital admission as well as potentially harmful and expensive treatments. Although the number of studies and patients are limited, absence of FBM appears to afford the highest diagnostic accuracy. Future studies should assess the diagnostic accuracy of the concomitant use of FBM and TVS CL measurement for the short-term prediction of preterm birth in symptomatic patients. ■

REFERENCES

1. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012;379:2162-72.
2. Beck S, Wojdyla D, Say L, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. *Bull World Health Organ* 2010;88:31-8.
3. Lawn JE, Kinney MV, Black RE, et al. Newborn survival: a multi-country analysis of a decade of change. *Health Policy Plan* 2012;3: 6-28.
4. Martin JA, Hamilton BE, Ventura SJ, Osterman MJK, Wilson EC, Matthews MS. Births: final data 2010. *Natl Vital Stat Rep* 2012;61.
5. Russell RB, Green NS, Steiner CA, et al. Cost of hospitalization for preterm and low birth weight infants in the United States. *Pediatrics* 2007;120:1-9.
6. Scott CL, Chavez GF, Atrash HK, Taylor DJ, Shah RS, Rowley D. Hospitalizations for severe complications of pregnancy, 1987-1992. *Obstet Gynecol* 1997;90:225-9.

7. Carbonne R. Is it possible to improve diagnostic and prognostic criteria of preterm labour? *Eur J Obstet Reprod Biol* 2004;117S:S6-9.

8. Sanchez-Ramos L, Delke I, Zamora J, Kaunitz AM. Fetal fibronectin as a short-term predictor of preterm birth in symptomatic patients: a meta-analysis. *Obstet Gynecol* 2009;114:631-40.

9. Timor-Tritsch IE, Boozarjomehri F, Masakowski Y, Monteagudo A, Chao CR. Can a «snapshot» sagittal view of the cervix by transvaginal ultrasonography predict active preterm labor? *Am J Obstet Gynecol* 1996;174:990-5.

10. Lukes AS, Thorp JM Jr, Eucker B, Pahel-Short L. Predictors of positivity for fetal fibronectin in patients with symptoms of preterm labor. *Am J Obstet Gynecol* 1997;176:639-41.

11. Ferguson JE, Dyson DC, Schutz T, Stevenson DK. A comparison of tocolysis with nifedipine or ritodrine: analysis of efficacy and maternal, fetal, and neonatal outcome. *Am J Obstet Gynecol* 1990;163:105-11.

12. Ferguson JE, Dyson DC, Holbrook RH, Schutz T, Stevenson DK. Cardiovascular and metabolic effects associated with nifedipine and ritodrine tocolysis. *Am J Obstet Gynecol* 1989;161:788-95.

13. King JG, Grant A, Keirse MJNC, Chalmers I. Beta-mimetics in preterm labour: an overview of the randomized controlled trials. *Br J Obstet Gynaecol* 1988;95:211-22.

14. Kemp VH, Hatmaker DD. Stress and social support in high-risk pregnancy. *Res Nursing Health* 1989;12:33106.

15. Merkatz R. Prolonged hospitalization of pregnant women: the effects on the family. *Birth Fam J* 1978;5:204-6.

16. Honest H, Bachmann LM, Sengupta R, Gupta JK, Kleijnen J, Khan KS. Accuracy of absence of fetal breathing movements in predicting preterm birth: a systematic review. *Ultrasound Obstet Gynecol* 2004;24:94-100.

17. Sotiriadis A, Papatheodorou S, Kawadias A, Makrydimas G. Transvaginal cervical length measurement for prediction of preterm birth in women with threatened preterm labor: a meta-analysis. *Ultrasound Obstet Gynecol* 2010;35:54-64.

18. Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: no good test for predicting a spontaneous preterm birth. *Curr Opin* 2012;24:422-33.

19. Honest H, Bachmann LM, Coomarasamy A, Gupta JK, Kleijnen J. Accuracy of cervical transvaginal sonography in predicting preterm birth: a systematic review. *Ultrasound Obstet Gynecol* 2003;22:305-22.

20. Honest H, Forbes CA, Duree KH, et al. Screening to prevent spontaneous preterm birth: systematic review of accuracy and effectiveness literature with economic modeling. *Health Technol Assess* 2009;43:1-627.

21. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283: 2008-12.

22. Devillé WL, Buntinx F, Bouter LM, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002;2:9.
23. Bossuyt PM, Reitsma JB, Bruns JB, et al. Toward complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 2003;49:1-6.
24. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;8:1-234.
25. Harbord RM, Whiting P, Sterne JA, et al. An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary. *J Clin Epidemiol* 2008;61:1095-103.
26. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008;149:889-97.
27. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-58.
28. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.
29. Lijmer JG, Bossuyt PM, Hesiterkamp SM. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 2002;21:1525-37.
30. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003;56:1129-35.
31. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy. *J Clin Epidemiol* 2005;58:1331-2.
32. O'Brien JM, Mercer B, Pitis DW, Sibai BM. A comparison of fetal fibronectin, prolactin and cervical examination for the prediction of preterm delivery in symptomatic patients. *Am J Obstet Gynecol* 1995;173:S406.
33. Iams JD, Casal D, McGregor JA, et al. Fetal fibronectin improves the accuracy of diagnosis of preterm labor. *Am J Obstet Gynecol* 1995;173:141-5.
34. Malak TM, Sizmur F, Bell SC, Taylor DJ. Fetal fibronectin in cervicovaginal secretions as a predictor of preterm birth. *BJOG* 1996;103:648-53.
35. Bartnicki J, Casal D, Kreaden US, Saling E, Better K. Fetal fibronectin in vaginal specimens predicts preterm delivery and very-low-birth-weight infants. *Am J Obstet Gynecol* 1995;174:971-4.
36. Senden IPM, Owen P. Comparison of cervical assessment, fetal fibronectin and fetal breathing in the diagnosis of preterm labour. *Clin Exp Obstet Gyn* 1996;23:5-9.
37. Benattar C, Taieb J, Fernandez H, Lindendaum A, Frydman R, Ville Y. Rapid fetal fibronectin swab-test in preterm labor patients treated by betamimetics. *Eur J Obstet Gynecol Reprod Biol* 1997;72:131-5.
38. Peaceman AM, Andrews WW, Thorp JM, et al. Fetal fibronectin as a predictor of preterm birth in patients with symptoms: a multicenter trial. *Am J Obstet Gynecol* 1997;177:13-8.
39. McKenna DS, Chung K, Iams JD. Effect of digital cervical examination on the expression of fetal fibronectin. *J Reprod Med* 1999;44:796-800.
40. Lopez RL, Francis JA, Garite TJ, DUBYAK JM. Fetal fibronectin detection as a predictor of preterm birth in actual clinical practice. *Am J Obstet Gynecol* 2000;182:1103-6.
41. LaShay N, Gilson G, Joffe G, Qualls C, Curet L. Will cervicovaginal interleukin-6 combined with fetal fibronectin testing improve the prediction of preterm delivery? *J Matern Fetal Med* 2000;9:336-41.
42. Giles W, Bisits A, Knox M, Madsen G, Smith R. The effect of fetal fibronectin testing on admissions to a tertiary maternal-fetal medicine unit and cost savings. *Am J Obstet Gynecol* 2000;182:439-42.
43. Closset E, Dufour P, Coeugnet C, Subtil D, Valat AS, Puech F. Value of fetal fibronectin research for predicting premature delivery [in French]. *Gynecol Obstet Fertil* 2001;29:808-13.
44. Ray D, Dyson D. Fetal fibronectin as a predictor of preterm delivery in symptomatic patients. *Am J Obstet Gynecol* 2001;184:S37.
45. Coleman MA, Keelan JA, McCowan LM, Townend KM, Mitchell MD. Predicting preterm delivery: comparison of cervicovaginal interleukin (IL)-1 β , IL-6 and IL-8 with fetal fibronectin and cervical dilatation. *Eur J Obstet Gynecol Reprod Biol* 2001;95:154-8.
46. Rinehart BK, Torrone DA, Isler CM, Barilleaux PS, Bufkin L, Morrison JC. Pregnancy outcome in women with preterm labor symptoms without cervical change. *Am J Obstet Gynecol* 2001;184:1004-7.
47. Luzzi V, Hankins K, Gronowski AM. Accuracy of the rapid fetal fibronectin TLI system in predicting preterm delivery. *Clin Chem* 2003;49:501-2.
48. Sakai M, Sasaki Y, Yamagishi N, Tanebe K, Yoneda S, Saito. The preterm labor index and fetal fibronectin for prediction of preterm delivery with intact membranes. *Obstet Gynecol* 2003;101:123-8.
49. Foxman EF, Jarolim P. Use of the fetal fibronectin test in decisions to admit to hospital for preterm labor. *Clin Chem* 2004;50:663-5.
50. Tekesin I, Marek S, Hellmeyer L, Reitz D, Schmidt S. Assessment of rapid fetal fibronectin in predicting preterm delivery. *Obstet Gynecol* 2005;105:280-4.
51. Gomez R, Romero R, Medina L, et al. Cervicovaginal fibronectin improves the prediction of preterm delivery based on sonographic cervical length in patients with preterm uterine contractions and intact membranes. *Am J Obstet Gynecol* 2005;192:350-9.
52. Schmitz T, Maillard F, Bessard-Bacquaert S, et al. Selective use of fetal fibronectin detection after cervical length measurement to predict spontaneous preterm delivery in women with preterm labor. *Am J Obstet Gynecol* 2006;194:138-43.
53. Skoll A, St Louis P, Amiri N, Delisle MF, Lalji S. The evaluation of the fetal fibronectin test for prediction of preterm delivery in symptomatic patients. *J Obstet Gynaecol Can* 2006;28:206-13.
54. Tsoi E, Akmal S, Geerts L, Jeffery B, Nicolaidis KH. Sonographic measurement of cervical length and fetal fibronectin testing in threatened preterm labor. *Ultrasound Obstet Gynecol* 2006;27:368-72.
55. Incerti M, Ghidini A, Korke V, Pezzullo JC. Performance of cervicovaginal fetal fibronectin in a community hospital setting. *Arch Gynecol Obstet* 2007;275:347-51.
56. Mateus J, Pereira L, Baxter J, Berghella V, Tolosa J. Effectiveness of fetal fibronectin testing compared with digital cervical assessment of women with preterm contractions. *Am J Perinatol* 2007;24:381-5.
57. Ting HS, Chin PS, Yeo GS, Kwek K. Comparison of bedside test kits for prediction of preterm delivery: phosphorylated insulin-like growth factor binding protein-1 (pIGFBP-1) test and fetal fibronectin test. *Ann Acad Med Singapore* 2007;36:399-402.
58. Woodworth A, Moore J, G'Sell C, et al. Diagnostic accuracy of cervicovaginal interleukin-6 and interleukin-6:albumin ratio as markers of preterm delivery. *Clin Chem* 2007;53:1534-40.
59. Tanir HM, Sener T, Yildiz Z. Cervicovaginal fetal fibronectin (FFN) for prediction of preterm delivery in symptomatic cases: a prospective study. *Clin Exp Obstet Gynecol* 2008;35:61-4.
60. Eroglu D, Yanik F, Oktem M, Zeyneloglu HB, Kuscu E. Prediction of preterm delivery among women with threatened preterm labor. *Gynecol Obstet Invest* 2007;64:109-16.
61. Peláez LM, Fox NS, Chasen ST. Negative fetal fibronectin: who is still treating for threatened preterm labor and does it help? *J Perinat Med* 2008;36:202-5.
62. Sunagawa S, Takagi K, Ono K, Miyachi K, Kikuchi A. Comparison of biochemical markers and cervical length for predicting preterm delivery. *J Obstet Gynaecol Res* 2008;34:812-9.
63. Wilsms FF, van Stralen G, Porath MM, et al. Predicating imminent preterm labour based on a determination of foetal fibronectin in a vaginal smear. *Ned Tijdschr Geneesk* 2009;153:B398.
64. Díaz J, Chedraui P, Hidalgo L, Medina M. The clinical utility of fetal fibronectin in the prediction of pre-term birth in a low socio-economic setting hospital in Ecuador. *J Matern Fetal Neonatal Med* 2009;22:89-93.
65. Deplagne C, Maurice-Tison S, Coatleven F, Vandebossche F, Horovitz J. Sequential use of cervical length measurement before fetal fibronectin detection to predict spontaneous preterm delivery in women with preterm labor. *J Gynecol Obstet Biol Reprod (Paris)* 2010;39:575-83.
66. López Farfán JA, Sánchez Tovar HB, Gutiérrez de Anda Mdel R, Gámez Guevara C. Fetal fibronectin and cervical length as early

predictors of preterm labor. *Ginecol Obstet Mex* 2011;79:337-43.

67. Burwick R, Zork N, Lee G, Ross M. Cervical assessment of cervical length compared to fetal fibronectin in the prediction of preterm delivery in women with threatened preterm labor. *J Matern Fetal Neonatal Med* 2011;24:127-31.

68. Riboni F, Vitulo A, Dell'avano M, Plebani M, Battagliarin G, Paternoster D. Biochemical markers predicting pre-term delivery in symptomatic patients: phosphorylated insulin-like growth factor binding protein-1 and fetal fibronectin. *Arch Gynecol Obstet* 2011;284:1325-9.

69. Van Baaren GJ, Vis J, Wilms F, et al. The accuracy of fetal fibronectin and cervical length in women with signs and symptoms of preterm labor before 34 weeks: a nationwide cohort study in The Netherlands (APOSTEL1). *Am J Obstet Gynecol* 2013;208: S8[Abstract #11].

70. Castle BM, Turnbull AC. The presence or absence of fetal breathing movements predicts the outcome of preterm labour. *Lancet* 1983;2:471-3.

71. Boylan P, O'Donovan P, Owens OJ. Fetal breathing movements and the diagnosis of labor: a prospective analysis of 100 cases. *Obstet Gynecol* 1985 Oct;66:517-20.

72. Jaschevatzky O, Ellenbogen A, Anderman S, Frisch L, Noy Y, Grunstein S. The predictive value of fetal breathing movements in the outcome of premature labour. *Br J Obstet Gynaecol* 1986;93:1256-8.

73. Agustsson P, Patel NB. The predictive value of fetal breathing movements in the diagnosis of preterm labour. *Br J Obstet Gynaecol* 1987;94: 860-3.

74. Besinger RE, Compton AA, Hayashi RH. The presence of absence of fetal breathing movements as a predictor of outcome in preterm labor. *Am J Obstet Gynecol* 1987;157: 753-7.

75. Schreyer P, Caspi E, Natan NB, Tal E, Weinraub Z. The predictive value of fetal breathing movement and Bishop score in the diagnosis of "true" preterm labor. *Am J Obstet Gynecol* 1989;161:886-9.

76. Kanaan CM, O'Grady JP, Beille JC. Effect of maternal carbon dioxide inhalation on human fetal breathing movements in term and preterm labor. *Obstet Gynecol* 1991;78:9-13.

77. Devoe LD, Youssef AE, Croom CS, Watson J. Can fetal biophysical observations anticipate outcome in preterm labor or preterm rupture of membranes? *Obstet Gynecol* 1994;84:432-8.

78. Markwitz W, Ropacka M, Breborowicz GH. Evaluation of fetal breathing movements in prognosis of preterm labor. *Ginekol Pol* 2001;72:55-60.

79. Tsoi E, Fuchs IB, Rane S, Geerts L, Nicolaidis KH. Sonographic measurement of cervical length in threatened preterm labor in singleton pregnancies with intact membranes. *Ultrasound Obstet Gynecol* 2005;25:353-6.

80. Gomez R, Romero R, Nien JK, et al. A short cervix in women with preterm labor and intact membranes: a risk factor for microbial invasion of the amniotic cavity. *Am J Obstet Gynecol* 2005;192:678-89.

81. Botsis D, Papagianni V, Vitoratos N, Makrakis E, Aravantinos L, Creatsas G. Prediction of preterm delivery by sonographic estimation of cervical length. *Biol Neonate* 2005;88:42-5.

82. Botsis D, Makrakis E, Papagianni V, et al. The value of cervical length and plasma proMMP-9 levels for the prediction of preterm delivery in pregnant women presenting with threatened preterm labor. *Eur J Obstet Gynecol Reprod Biol* 2006;128:108-12.

83. Holst RM, Jacobsson B, Hagberg H, Wennerholm UB. Cervical length in women in preterm labor with intact membranes: relationship to intra-amniotic inflammation/microbial invasion, cervical inflammation and preterm delivery. *Ultrasound Obstet Gynecol* 2006;28: 768-74.

84. Gramellini D, Fieni S, Kaihura C, Modena AB. Cervical length as a predictor of preterm delivery: gestational age-related percentiles vs fixed cut-offs. *Acta Biomed* 2007;78:220-4.

85. Palacio M, Sanin-Blair J, Sánchez M, et al. The use of a variable cut-off value of cervical length in women admitted for preterm labor before and after 32 weeks. *Ultrasound Obstet Gynecol* 2007;29:421-6.

86. Boudhraa K, Rahouej H, Gara MF. Transvaginal ultrasound of the cervix in the estimation of severity of premature labour. *Tunis Med* 2008;86:745-8.

87. Schmitz T, Kayem G, Maillard F, Lebret MT, Cabrol D, Goffinet F. Selective use of sonographic cervical length measurement for predicting imminent preterm delivery in women with preterm labor and intact membranes. *Ultrasound Obstet Gynecol* 2008;31:421-6.

88. Danti L, Prefumo F, Lojaco A, Corini S, Testori A, Frusca T. The combination of short cervical length and phIGFBP-1 in the prediction of preterm delivery in symptomatic women. *J Matern Fetal Neonatal Med* 2011;24:1262-6.

89. Wulff CB, Ekelund CK, Hedegaard M, Tabor A. Can a 15-mm cervical length cutoff discriminate between low and high risk of preterm delivery in women with threatened preterm labor? *Fetal Diagn Ther* 2011;29:216-23.

90. Demirci O, Ünal A, Demirci E, et al. Sonographic measurement of cervical length and risk of preterm delivery. *J Obstet Gynaecol Res* 2011;37:809-14.

91. Adhikari K, Bagga R, Suri V, Takhtani M. Cervical length compared to Bishop's score for prediction of pre-term birth in women with preterm labour. *J Obstet Gynaecol* 2011;31:213-6.

92. Giannella L, Beraldi R, Giulini S, Cerami LB, Mfuta K, Facchinetti F. Nitric oxide metabolite levels and assessment of cervical length in the prediction of preterm delivery among women

undergoing symptomatic preterm labor. *Int J Gynaecol Obstet* 2012;116:223-7.

93. Perales-Puchalt A, Brik M, Diago VJ, Perales A. The negative predictive value of cervical interleukin-6 for the risk assessment of preterm birth. *J Matern Fetal Neonatal Med* 2013;26:1278-81.

94. Geersing GJ, Janssen KJM, Oudega R, et al. Excluding venous thromboembolism using point of care D-dimer tests in outpatients: a diagnostic meta-analysis. *BMJ* 2009;339:b2990.

95. Crane JM, Van den Hof MC, Armson BA, Liston R. Transvaginal ultrasound in the prediction of preterm delivery: singleton and twin gestations. *Obstet Gynecol* 1997;90:357-63.

96. Murakawa H, Utumi T, Hasegawa I, Tanaka K, Fuzimari R. Evaluation of threatened preterm delivery by transvaginal ultrasonographic measurement of cervical length. *Obstet Gynecol* 1993;82:829-32.

97. Goldenberg RL, Mercer BM, Meis PJ, Copper RL, Das A, McNellis D. The preterm prediction study: fetal fibronectin testing and spontaneous preterm birth. *Obstet Gynecol* 1996;87:643-8.

98. Leitch H, Brunbauer M, Kaider A, Egarter C, Husslein P. Cervical length and dilation of the internal cervical os detected by vaginal ultrasonography as markers for preterm delivery: a systematic review. *Am J Obstet Gynecol* 1999;181:1465-72.

99. Gyetuai K, Hannah M, Hodnett E, Ohlsson A. Tocolytics for preterm labor: a systematic review. *Obstet Gynecol* 1999;94:869-77.

100. King JF, Grant A, Kerse M, Chalmers I. Betamimetics in preterm labor: an overview of the randomized controlled trials. *Br J Obstet Gynaecol* 1998;95:211-22.

101. Khan KS, Khan SF, Nwosu CR, Amott N, Chien PF. Misleading authors' inferences in obstetric diagnostic test literature. *Am J Obstet Gynecol* 1999;181:112-5.

102. Abbott DS, Radford SK, Seed PT, Tribe RM, Shennan AH. Evaluation of a quantitative fetal fibronectin test for spontaneous preterm birth in symptomatic women. *Am J Obstet Gynecol* 2013;208:122.e1-6.

103. Alfirevic Z, Allen-Coward H, Molina F, Vinuesa CP, Nicolaidis K. Targeted therapy for threatened preterm labor based on sonographic measurement of the cervical length: a randomized controlled trial. *Ultrasound Obstet Gynecol* 2007;29:47-50.

104. DeFranco EA, Lewis DF, Odibo AO. Improving the screening accuracy for preterm labor: is the combination of fetal fibronectin and cervical length in symptomatic patients a useful predictor of preterm birth? A systematic review. *Am J Obstet Gynecol* 2013;208:233.e1-6.

105. Egger M, Davey Smith G, Schneider M. Systematic reviews of observational studies. In: Egger M, Davey Smith G, Altman D, eds. *Systematic reviews in healthcare: meta-analyses in context*. London: BMJ Publishing group; 2001:211-27.