Ambrogio P. Londero*, Ralf Schmitz, Serena Bertozzi, Lorenza Driul and Arrigo Fruscalzo*

Diagnostic accuracy of cervical elastography in predicting labor induction success: a systematic review and meta-analysis

Abstract

Aim: To determine the accuracy of cervical elastography in predicting labor induction success.

Materials and methods: A systematic search, review, and meta-analysis of observational studies published in English language between January 2000 and October 2014 was performed. It included studies considering cervix sonoelastography as the index test and successful labor or vaginal delivery as the reference standard. As cervix length and Bishop score were considered comparator tests, the quality of the included studies was assessed using quality assessment tool for diagnostic accuracy studies (QUADAS) tool.

Results: A total of four studies assessing 323 women before medical induction of labor were included. Cervical elastography, cervical length, and Bishop score showed a diagnostic odds ratio (DOR) with 95% confidence interval (CI) for successful labor prediction of 3.50 (1.93–6.35), 3.35 (1.94–5.77), and 1.45 (0.33–6.41), respectively. In addition, cervical elastography, cervical length, and Bishop score showed a DOR with 95% CI for successful vaginal delivery prediction of 5.24 (3.23–8.50), 4.94 (2.72–8.98), and 4.62 (0.69–30.94), respectively. Considering the summary of receiver operating characteristic curves we show that cervical elastography or length are similarly reliable, and both are more reliable to predict successful labor than the Bishop score. Two studies were excluded because it was not possible to retrieve data for the meta-analysis. Among the excluded studies, one found no significant contribution from elastography for prediction of successful labor induction.

Conclusions: Even though there is a limited number of studies included and the heterogeneity of the methods used, cervical elastography seems to be a promising tool for predicting successful labor induction and vaginal delivery in women treated by medical induction of labor.

Keywords: Cervix; labor induction; sensitivity and specificity; tissue elastography; ultrasound; ultrasound tissue characterization.

DOI 10.1515/jpm-2015-0035
Received January 21, 2015. Accepted April 27, 2015.

***Corresponding authors: Dr. Ambrogio P. Londero,** Unit of Obstetrics and Gynecology, S. Polo Hospital, 34074 Monfalcone (GO), Italy; and DSMB, DISM, University of Udine, 33100 Udine, Italy, E-mail: ambrogio.londero@gmail.com; and **Arrigo Fruscalzo:** Obstetrics and Gynecology, St. Franziskus Hospital Münster, 48145 Münster, Germany; and Department of Obstetrics and Gynecology, University Hospital of Münster, 48149 Münster, Germany, Tel.: +49/02 51/9 35-39 17, Fax: +49/02 51/9 35-40 72, E-mail: fruscal@libero.it
Ralf Schmitz: Department of Obstetrics and Gynecology, University Hospital of Münster, 48149 Münster, Germany
Serena Bertozzi and Lorenza Driul: DSMB, DISM, University of Udine, 33100 Udine, Italy

Introduction

Labor induction is one of the most common interventions in clinical obstetrics, occurring in around 22% of all gravid women in the United States of America [1]. Whether medically indicated or elective for post-term pregnancies, it is associated with an increased risk for cesarean delivery, particularly in nulliparous women [2]. Predicting the success and duration of an induced labor is still a mostly unresolved goal for obstetricians. Despite many scoring systems having been proposed [3–5], the Bishop score is the most commonly used yet. The Bishop score is a 10-point scoring system achieved by assessing the following five components on vaginal examination: cervical dilation, cervical effacement, cervical consistency, cervical position, and fetal station [3]. A recent meta-analysis has shown that despite the Bishop score having been proven to be suitable for predicting successful vaginal delivery, no association was found with induction to active phase time interval [6].

Cervical length measurement was proposed as a predictor of successful labor induction and vaginal delivery in the literature. The effort was to replace the Bishop score with an objective, more reliable, and easy-to-perform diagnostic method supported by transvaginal ultrasound. Nonetheless, a recent meta-analysis revealed that even if assessment pooling results from multiple clinical trials were weakened by the wide variability of the used study protocols, subgroup analyses demonstrated that cervical length did not accurately predict any single specific outcome. Furthermore, no significant differences were found between the diagnostic accuracy of cervical length compared to Bishop score [7].

The topic of cervical stiffness estimation performed using ultrasound tools was already addressed in the past. Among other tools [8], quantitative ultrasonic tissue characterization and the ultrasonic attenuation estimation were proposed. Nonetheless, a high variability between subjects and the inability to standardize these tools was evidenced [9, 10]. The recent introduction of cervical elastography has boosted the interest in application of this new tool in the field of obstetrics [11]. This new ultrasound tool aims to allow an objective evaluation of the cervical stiffness, thus potentially replacing the current subjective evaluation by digital palpation. Several approaches have been proposed with encouraging preliminary results, and now clinical studies evaluating the effectiveness of cervical elastography in predicting preterm delivery [12–16] and a successful labor induction are rapidly increasing [8, 17–24]. We performed this meta-analysis to assess the accuracy of cervix elastography in predicting labor induction success.

Materials and methods

This study is a systematic review and meta-analysis of studies assessing the diagnostic test accuracy of cervix ultrasound tissue characterization in predicting labor induction success and vaginal delivery among women undergoing labor induction.

Search strategy for review

A systematic search of Google Scholar (Google Inc. Mountain View, CA, USA), MEDLINE (U.S. National Library of Medicine, Bethesda, MD, USA), Scopus (Elsevier BV, Amsterdam, the Netherlands), Ovid (Wolters Kluwer, Alphen aan den Rijn, the Netherlands), and Cochrane database (The Cochrane Collaboration, John Wiley & Sons, Inc.) was performed for studies published from January 2000 to December 2014. The search strategy was developed using the following terms "cervix elastography", "cervix strain", "cervix consistency", and "cervix quantitative ultrasound tissue characterization". All retrieved items were exported in a bibliography manager.

Inclusion and exclusion criteria

All studies that assessed the diagnostic accuracy of cervix ultrasound tissue characterization in predicting labor induction success and vaginal delivery among women undergoing labor induction were eligible for inclusion. Studies had to include women undergoing labor induction. No age restriction for study participants was considered. Both prospective and retrospective data collections were considered and no randomized trials have been found. In this meta-analysis, we included studies reporting the numbers of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) when analyzing the diagnostic accuracy of cervix ultrasound tissue characterization in predicting labor induction success or vaginal delivery. The cervix ultrasound tissue characterization could be performed alone or in combination with other diagnostic modalities (cervical length or bishop score). We gathered from full text articles study period and geographic location in order to avoid any possible population overlap. In case of two or more studies presenting possible data overlap, only the study of better quality or with more detailed data was included. Non-English written articles, studies considering <20 patients, studies about nonhuman subjects, or studies with unknown number of TP, FP, FN, or TN were specific exclusion criteria. In addition, editorials, review articles, and letters to editors that did not report original data were excluded.

Index and comparator tests

The cervical elastography was considered as the index test. Different techniques have been described in the literature mainly due to the different approaches adopted by commercial ultrasound scanner manufacturers. Despite the fact that it would be desirable for the various ultrasound manufacturers to adopt more uniform methods to assess the tissue stiffness, for our purpose all studies that used an ultrasound-based method to assess cervical tissue characterization were considered eligible. The main methods used to assess cervical tissue stiffness in obstetrics were the following: Elastoscan™ (Samsung Medison) – a cervical elastography tool integrated in some Accuvix scanners where the amount of tissue movements were represented on a color map allowing visual semiquantitative evaluation of tissue stiffness or software-aided score calculation by image analysis performed with the analySIS TS automatic program (Olympus Optical Co, Ltd, Tokyo, Japan) [21, 22]. Another method used was tissue Doppler imaging (TDI) (Toshiba Medical Systems, Tokyo, Japan) that is a Doppler-based tool for the imaging and estimation of tissue strain [14, 19, 25–27]. TDI allows for the tracking of tissue movement, whereas the TDI-Q (Q-Quantification) software (Toshiba Medical Systems, Tokyo, Japan) facilitates an estimation of tissue stiffness. Even though this software was designed for cardiac imaging, its use has been largely established in a lot of other medical fields (hepatology, breast disease, endocrinology, or surgery), working with extrinsic freehand movement of tissue compression exerted through the ultrasound probe [19]. Another method used to estimate the approximate

Young's modulus of the anterior cervical lip is by using a reference cap applied on the end of the transvaginal transducer during elastography and using a Voluson E8 Expert scanner with the BT13 and H48681GB software (GE Healthcare Austria, Zipf, Austria) [20]. Although cervical length or Bishop score were considered as comparator tests in the cervical elastography literature, in this study no comparator test was considered for eligibility.

Target conditions and reference standard

In this study, the index test was considered to predict successful labor induction or vaginal delivery as target conditions. However, in the literature there are different reference standards to define successful labor induction and vaginal delivery. For the purpose of this work, we considered as successful labor induction or vaginal delivery all the definitions considered by the selected studies.

Selection of studies and data extraction

After study retrieval from databases we checked for duplicates and examined all titles and abstracts to select all the articles that referred to the accuracy of cervix ultrasound tissue characterization in predicting labor induction success or vaginal delivery among women undergoing labor induction. Then, we analyzed the full articles of the selected items. Finally, reference lists and citations from full articles and previous review publications were searched to identify other additional pertinent articles. Moreover, three reviewers independently selected the studies and extracted data from the full-text articles. In case of discrepancies, a joint reevaluation of the original article was performed to address them. The three reviewers used a data-extraction form to gather from the articles data on patient characteristics (sample size, mean age), technical aspects of ultrasound elastography (number of performers and assessors, compression method, ultrasound scanners used), and scores used to determine test prediction accuracy for successful labor induction and vaginal delivery. For extraction of diagnostic accuracy data, the number of TP, FP, FN, and TN were extracted from the text or the tables. If the primary studies did not report the exact number of TP, FP, FN, and TN, the expected frequency was calculated using the total number of positives and negatives and the reported sensitivity and specificity in each study.

Assessment of methodological quality

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool was used to assess the quality of published studies [28, 29]. In this study we used the modified QUADAS tool that included 11 items instead of 14 [28, 29]. In addition, the quality of the studies was assessed considering a further nine items introduced by the Cochrane Collaboration [28]. Every item consisted of a question with three possible answers (yes, no, or unclear) [28]. A judgment of "yes" refers to the optimal methodological characteristic; a judgment of "no" refers to the less than optimal methodological characteristic; and a judgment of "unclear" refers to the category of uncertainty.

Data analysis

Data was analyzed by R (version 3.1.1, R Foundation for Statistical Computing, Vienna, Austria) and mada package version 0.5.5. Pooled estimates for positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) with the corresponding 95% CIs were used to examine the accuracy of different cervix ultrasound tissue characterization score thresholds for predicting labor induction success and vaginal delivery. Where appropriate, the fixed- and the random-effect model to calculate the pooled estimate were applied. The Cochran's Q test was used to assess heterogeneity among studies. The inconsistency (I^2) index was used to detect the percentage of variability due to heterogeneity rather than sampling errors. As previously described, an I^2 index value >50% and, a Q statistic P value <0.10 were considered statistically significant signs for heterogeneity [30]. In addition, we plotted the summary receiver operating characteristic (sROC) curves. To assess the presence of publication bias, Egger's test (considering P value <0.05 statistically significant) and funnel plots [with standard error (SE) on the vertical axis and the logarithm of the DOR on the horizontal axis] for asymmetry were used [31, 32]. The Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines for accurate performing meta-analysis of observational studies [33] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines checklist [34] were considered to prepare this meta-analysis.

Results

Search results

The adopted search strategy found 240 studies out of which 234 were excluded by reviewing abstracts, titles, and MEDLINE meta-data. After reviewing the full-text of the selected six articles, we excluded two further articles because it was not possible to extract the numbers of TP, FP, FN, and TN for an established score threshold [18, 23]. Finally, four studies were included for meta-analysis [19–22]. In Figure 1, the study selection flow diagram is shown.

Characteristics of the studies

The included studies, published between 2013 and 2014, considered 323 women enrolled for labor induction at term of pregnancy [19–22]. In Table 1, the characteristics of the included and excluded studies are shown. Three studies were conducted in European countries [19, 20, 22] and one study in South Korea [21]. All studies were prospective and included term pregnancies. Labor induction was made by different protocols. Hwang et al. induced labor by oxytocin with or without amniorexi [21], Fruscalzo et al. and Hee et al. used a protocol based on prostaglandin usage

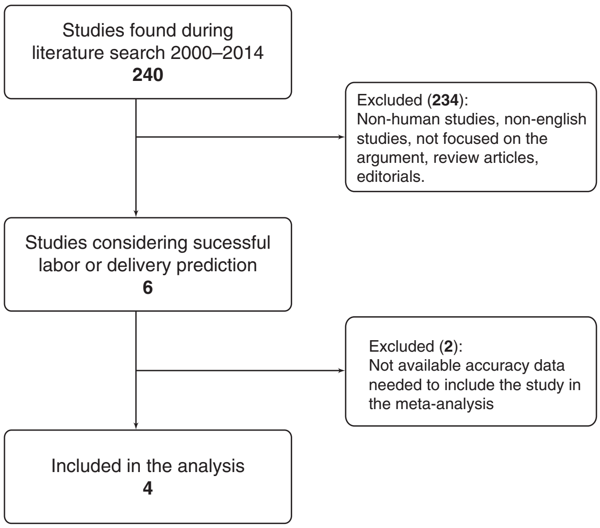

Figure 1: Flow-chart of the literature search and selection.

[19, 20], while Muscatello et al. did not specify the method of labor induction [22]. The most common indication for labor induction in all the studies was prolonged pregnancy with a median prevalence of 38% (range 35%–60%) followed by pregnancy related hypertensive disorders with a median prevalence of 11% (range 2%–28%). Then, among other indications listed were glucose metabolism alterations, fetal growth restriction, fetal macrosomia, oligohydramnios, or maternal indication.

The methods to assess cervix ultrasound tissue characteristics were different in the included studies. However, the most common technology used in two of the four included studies was the one of Samsung-Medison® (Elastoscan™). Also concerning the reference standards there were significant differences in the included studies. Hwang et al. defined the successful induction of labor as the onset of active labor within 9 h and the successful delivery when it occurred within 24 h after labor induction [21]. Fruscalzo et al. took in consideration most of the common endpoints considered in the literature to define a successful labor induction, but for the purpose of this study we choose the closest to that of Hwang et al. even if in the original study these cut-offs were not significantly predicted by cervix ultrasound tissue characteristics (only the overall failure of labor induction was significantly predicted by ultrasound tissue characteristics) [19, 21]. Therefore, successful induction was considered the onset of active labor within 24 h and the successful delivery when it occurred vaginally within 36 h after labor induction [19]. Muscatello et al. considered endpoint vaginal delivery and Hee et al. considered the time of cervical dilatation [20, 22]. Other characteristics of the studies are presented in Table 1.

Quality assessment of the included studies and publication bias

Figure 2A illustrates the methodological quality of the included studies based on the modified QUADAS tool by the Cochrane Collaboration. The summary of the compliance of individual studies to these 11 items and of the risk of bias is shown in Figure 2B. Information for the additional nine items proposed by the Cochrane Collaboration are provided in Figure 2C. In all the included studies an acceptable reference standard independent of the index test was used and the sample of patients was representative. Index test results were blinded in all cases. In three of the four studies it was not clear if the reference standard was decided blinded to the results of cervix ultrasound tissue characterization (index test) or not [20–22]. No mention of uninterpretable results was given in the three studies [19, 21, 22]. In none of the included studies was the threshold of the index test established as *a priori*. Only one study mentioned the training of the operators stating that they were experienced sonographers [20]. None of the studies provided any statement about specific training with the ultrasound technique used to assess ultrasound cervical tissue characterization. Inter-observer and instrument variation were not reported in two articles [19, 22]. However, concerning the elastographic approach adopted by Fruscalzo et al., the intra- and inter-observer variability and instrument variation has already been evaluated in a separate study [27]. In another paper it was not clear if instrument variation was tested [21]. Two studies [19, 20] were reported to be free of conflicts of interest due to commercial funding, but in the other two studies [21, 22], there were no data about conflicts of interest due to the source of funding.

In addition, the presence of publication bias was examined by funnel plot and Egger's test (Figure 2C). In both analyses, no publication bias was observed. However, these results should be considered with caution, because our meta-analysis calculation included only four studies. In fact, current guidelines do not recommend testing for the funnel plot asymmetry or for the correlation test of funnel plot asymmetry in the analysis of a limited number of studies (<10) [35].

Main analysis

In Tables 2 and 3 we show the accuracy results of cervical elastography to predict successful labor induction and vaginal delivery. A total of six studies assessing 323 women before medical induction of labor were included

Table 1: Characteristics of included and excluded studies.

Author	Year study was done	Year study was published	Number of patients	Mean/median patient's age	Nulliparous	Gestational age (weeks)	Labor induction	Software	Reference standard (Successful induction of labor)
Included studies									
Hwang et al. 2013 [21]	2011	2013	145	31 (22–42)	100%	37–42 weeks (range)	Oxytocin/amniorexi	Elastoscan (Samsung-Medison) and analySIS TS automatic program (Olympus Optical Co)	Onset of active labor within 9 h or delivery within 24 h after labor induction
Fruscalzo et al. 2014 [27]	2010–2011	2014	77	29.7 (±3.5)	58%	39.7 (±1.5) weeks (mean)	Dinoproston	TDI and TDI-Q (Toshiba)	Time to labor – time to delivery – mg of prostaglandin usage – vaginal delivery
Muscatello et al. 2014 [22]	NA	2013	53	31.6 (21–44)	62%	38–41 weeks (range)	NA	Elastoscan (Samsung-Medison)	Successful delivery
Hee et al. 2014 [20]	NA	2014	48	31 (±4.6)	54%	284 days (±9.5) (mean)	Misoprostol	BT13 software/H48681GB software (GE)	Prolonged cervical dilatation time (>330 min)
Excluded studies									
Swiatkowska-Freund and Price 2011 [18]	NA	2011	29	27 (18–39)	0 (0–1) vs 0 (0–3)[a]	40 (33–42) vs 38 (36–41)[a], 39 weeks (mean)	Oxytocin	ElastoScan (Medison)	Uterine contractions and labor progress within 9 h
Pereira et al. 2014 [23]	April–October 2013.	2014	99	32.0 (29.0–35.3) vs 33.0 (30.0–35.5)[b]	41 (62.1) vs 31 (93.9)[b]	41.4 (39.3–41.6) vs 41.0 (38.7–41.6)[b]	Dinoprostone (10 mg) slow-release vaginal pessary or dinoprostone (3 mg) vaginal tablet or amniorexi, ±Oxytocin	ElastoScan (Accuvix XG, Samsung-Medison), and software "stiffmetool" (Samsung-Medison)	Vaginal delivery

[a]Failure vs success, [b]Vaginal delivery vs cesarean section; NA=not specified.

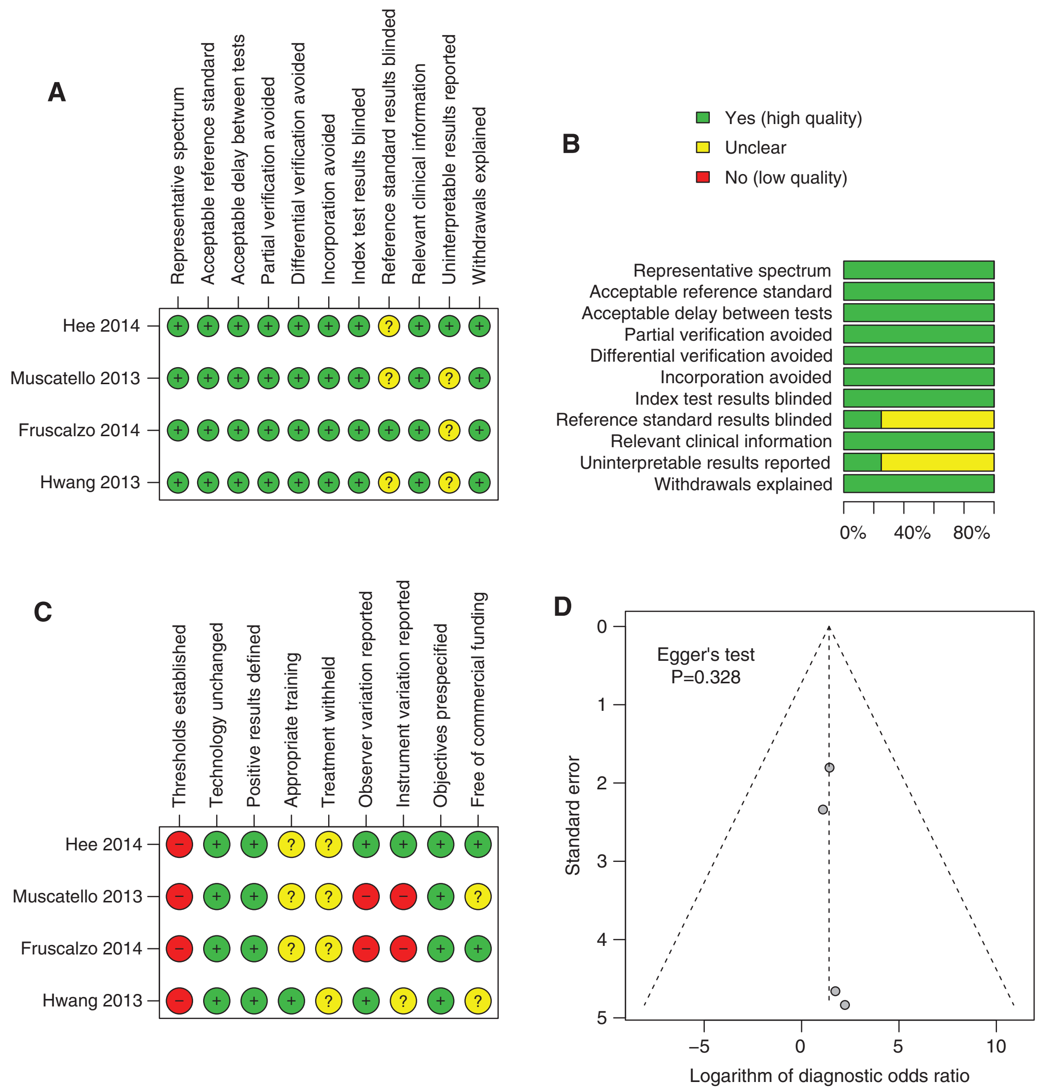

Figure 2: Quality assessment and publication bias. (A) Methodological quality summary about each risk-of-bias item for each study included in this meta-analysis [colors legend: green (+) as yes (high quality), yellow (?) as unclear, and red (–) as no (low quality)]. (B) Methodological quality graph shows each methodological quality item summary (in percentage of all studies included in this meta-analysis). (C) Methodological quality summary of additional quality items proposed by Cochrane Collaboration. (D) Funnel plot of diagnostic odds ratios for cervical ultrasound tissue characterization and Egger's test.

of which 270 were evaluated for successful labor and 137 had a successful labor induction. Furthermore, 275 were evaluated for successful vaginal delivery and 137 had a successful vaginal delivery.

Cervix ultrasound tissue characterization showed a DOR, sensitivity, and specificity for successful labor induction prediction of 3.50 (1.93–6.35), 71.1% (58.1%–81.4%), and 54.7% (48.2%–61.1%), respectively (Table 2). Moreover, cervix ultrasound tissue characterization showed a DOR, sensitivity, and specificity for successful vaginal delivery prediction of 5.24 (3.23–8.50), 68.2% (44.1%–85.3%), and 67.7% (79.8%–52.7%), respectively (Table 3). Therefore, cervix ultrasound tissue characterization seems a reliable method to predict successful labor induction and vaginal delivery. In Tables 2 and 3, we also reported the diagnostic accuracy of cervical length and Bishop score considering the selected articles. We found that cervix elastography had the highest DOR in comparison to cervical length and Bishop score for successful labor induction or vaginal delivery. In addition, we calculated summary receiver operating characteristic curves (sROC) (Figure 3). Then, considering sROC curves to predict successful labor, we showed that the summary estimates of cervical sonoelastography or cervical length

Table 2: Meta-analytic summary of diagnostic accuracy measures to predict successful labor for cervical ultrasound tissue characterization, cervical length, and Bishop score considering the included studies.

	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)
Labor success					
Cervix elastography					
Hwang et al. 2013 [21] – Elastographic index	70.8% (56.8%–81.8%)	50.5% (40.7%–60.3%)	1.431 (1.092–1.877)	0.577 (0.356–0.936)	2.479 (1.184–5.19)
Hwang et al. 2013 [21] – Cervical hard area	83.3% (70.4%–91.3%)	55.7% (45.8%–65.2%)	1.88 (1.455–2.429)	0.299 (0.155–0.578)	6.279 (2.662–14.813)
Fruscalzo et al. 2014 [27] – Cervix TS	55.6% (42.4%–68%)	60.9% (40.8%–77.8%)	1.42 (0.809–2.492)	0.73 (0.469–1.137)	1.944 (0.719–5.257)
Hee et al. 2014 [20] – Young's modulus	74.3% (57.9%–85.8%)	69.2% (42.4%–87.3%)	2.414 (1.044–5.583)	0.371 (0.19–0.726)	6.5 (1.602–26.375)
Pooled estimate	71.1% (58.1%–81.4%)	54.7% (48.2%–61.1%)	1.655 (1.384–1.980)	0.495 (0.334–0.733)	3.501 (1.930–6.349)
Cervix length					
Hwang et al. 2013 [21]	70.8% (56.8%–81.8%)	58.8% (48.8%–68%)	1.718 (1.274–2.316)	0.496 (0.31–0.795)	3.461 (1.648–7.269)
Fruscalzo et al. 2014 [27]	72.2% (59.1%–82.4%)	60.9% (40.8%–77.8%)	1.846 (1.08–3.154)	0.456 (0.266–0.784)	4.044 (1.447–11.301)
Hee et al. 2014 [20]	65.7% (49.2%–79.2%)	53.8% (29.1%–76.8%)	1.424 (0.755–2.684)	0.637 (0.322–1.258)	2.236 (0.613–8.161)
Pooled estimate	70% (61.8%–77.1%)	58.6% (50.1%–66.7%)	1.696 (1.332–2.159)	0.509 (0.371–0.697)	3.346 (1.939–5.775)
Bishop score					
Hwang et al. 2013 [21]	56.2% (42.3%–69.3%)	30.9% (22.6%–40.7%)	0.814 (0.614–1.081)	1.415 (0.913–2.191)	0.576 (0.282–1.176)
Fruscalzo et al. 2014 [27]	38.9% (27%–52.2%)	95.7% (79%–99.2%)	8.944 (1.278–62.606)	0.639 (0.508–0.804)	14 (1.754–111.767)
Hee et al. 2014 [20]	51.4% (35.6%–67%)	46.2% (23.2%–70.9%)	0.955 (0.526–1.736)	1.052 (0.534–2.075)	0.908 (0.253–3.251)
Pooled estimate	48.8% (36.2%–61.6%)	63.7% (16.1%–94.1%)	1.089 (0.564–2.103)	0.954 (0.541–1.685)	1.455 (0.330–6.408)

PLR=positive likelihood ratio, NLR=negative likelihood ratio, DOR=diagnostic odds ratio, CI=confidence interval, TS=tissue strain.

Table 3: Meta-analytic summary of diagnostic accuracy measures to predict successful vaginal delivery for cervical ultrasound tissue characterization, cervical length, and Bishop score considering the included studies.

	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)
Vaginal delivery success					
Cervix elastography					
Hwang et al. 2013 [21]– Elastographic index	76.5% (63.2%–86.0%)	56.4% (46.3%–66.0%)	1.75 (1.33–2.31)	0.42 (0.25–0.70)	4.20 (1.96–9.03)
Hwang et al. 2013 [21]– Cervical hard area	86.3% (74.3%–93.2%)	59.6% (49.5%–68.9%)	2.13 (1.63–2.79)	0.23 (0.11–0.47)	9.26 (3.77–22.73)
Fruscalzo et al. 2014 [27]– Cervix TS	35.3% (23.6%–49.0%)	84.6% (66.5%–93.8%)	2.29 (0.86–6.08)	0.76 (0.59–0.99)	3.00 (0.89–10.06)
Muscatello et al. 2014 [22]– Elastographic index	68.6% (52.0%–81.4%)	72.2% (49.1%–87.5%)	2.47 (1.13–5.37)	0.435 (0.25–0.78)	5.673 (1.62–19.88)
Pooled estimate	68.2% (44.1%–85.3%)	67.7% (79.8%–52.7%)	1.98 (1.656–2.37)	0.450 (0.27–0.75)	5.24 (3.23–8.50)
Cervix length					
Hwang et al. 2013 [21]	64.7% (51.0%–76.4%)	70.2% (60.3%–78.5%)	2.17 (1.50–3.15)	0.503 (0.34–0.75)	4.32 (2.09–8.92)
Fruscalzo et al. 2014 [27]	80.4% (67.5%–89%)	61.5% (42.5%–77.6%)	2.09 (1.26–3.46)	0.32 (0.17–0.6)	6.56 (2.30–18.74)
Pooled estimate	72.9% (54.5%–85.8%)	66.6% (54.4%–76.9%)	2.14 (1.59–2.89)	0.43 (0.28–0.66)	4.94 (2.72–8.98)
Bishop score					
Hwang et al. 2013 [21]	62.7% (49%–74.7%)	56.4% (46.3%–66%)	1.44 (1.05–1.97)	0.66 (0.44–0.98)	2.18(1.08–4.38)
Fruscalzo et al. 2014 [27]	39.2% (27%–52.9%)	96.2% (81.1%–99.3%)	10.196 (1.45–71.81)	0.63 (0.50–0.80)	16.13 (2.02–128.63)
Pooled estimate	51.0% (27.9%–73.7%)	83.7% (21.9%–98.9%)	2.992 (0.47–19.17)	0.64 (0.52–0.78)	4.62 (0.69–30.94)

PLR=positive likelihood ratio, NLR=negative likelihood ratio, DOR=diagnostic odds ratio, CI=confidence interval, TS=tissue strain.

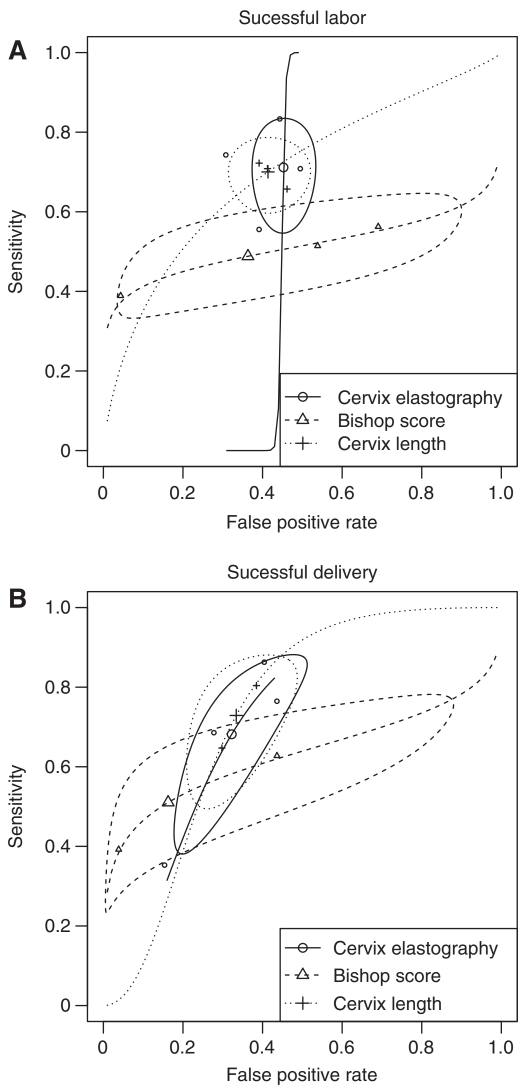

Figure 3: Summary receiver operating characteristic curves (sROC) considering the included studies. (A) Prediction of successful labor (AUC: cervical elastography 54.7%; Bishop score 50.9%; cervix length 69.7%). (B) Prediction of successful vaginal delivery (AUC: cervical elastography 72.5%; Bishop score 62.5%; cervix length 73.2%).
The summary estimates of both cervical elastography and cervical length are superior and well separated from the Bishop score in the prediction of a successful labor induction (Figure 3A), whereas their confidence regions were partially overlapping to the Bishop score in the prediction of a successful vaginal delivery (Figure 3B). Therefore, both cervix elastography and cervical length show to be a more reliable way than the Bishop score to predict a successful labor induction in the selected studies.

are well separated from Bishop score, though the confidence regions slightly overlap (Figure 3A). Therefore, it would be safe to conclude that cervix sonoelastography or length are a more reliable way to predict successful labor than Bishop score in the selected studies.

Risk of bias assessment

All the included studies were observational prospective studies and in most of the cases representing the initial experience with this innovative ultrasound technique. Each considered group adopted different techniques to assess cervical ultrasound tissue characteristics. This in part is due to the inhomogeneity of the ultrasound systems used (every manufacturer of ultrasound systems uses different approaches) and in part because there is still no standardized approach to assess cervical ultrasound tissue characteristics. Despite this inhomogeneity, the majority of published studies agree to pursue this interesting technique in the field of obstetrics.

Discussion

Summary of our results

This is the first meta-analysis dealing with the topic of clinical usefulness of cervical elastography during pregnancy. The accuracy of cervical elastography for prediction of a successful labor induction and vaginal delivery were evaluated and compared to the current gold standard, the Bishop score, and to the cervical length measured by transvaginal ultrasound. Four studies were included in the meta-analysis. Despite the different methods used for cervical elastography, results were consistent among different studies. Furthermore, results indicated an accuracy of cervical elastography in the prediction of successful labor induction and vaginal delivery overlapping to that of cervical length measurement and superior to that of Bishop score.

Endpoints of the studies

The two endpoints chosen in our meta-analysis, the labor induction success and vaginal delivery, are generally the two most relevant endpoints when evaluating labor induction success. Even if different definitions for labor induction success have been used, this is difficult to avoid, and reflects the difficulty of finding a reliable definition for labor induction success. Nonetheless, it appears to be probably the most interesting parameter to be evaluated, as vaginal delivery success can be influenced by a lot of other obstetrical factors not necessarily implicated in cervical dystocia, for example, intrapartum fetal distress.

The Bishop score

Despite its limitations, the Bishop score is to date the gold standard in the prediction of labor induction success. The strength of this method resides in the multiplicity of evaluated parameters, not limited to the cervix that allows to trace a global trait of the pregnancy course [3]. The major limitation includes the semi-quantitative and subjective evaluation, with the consequent difficulty of quantification and low reliability of evaluation. Furthermore, most literature describes the Bishop score as a useful parameter for predicting labor induction mainly in patients with favorable pelvic scores [36].

The transvaginal cervical length measurement

Cervical length also varies during a physiological pregnancy, depending above all on gestational week of scan, parity, and other maternal anthropometric factors [37, 38]. Methods to measure cervical length are currently debated. It has been proposed that cervical length can be measured by a straight line or by a traced line. With regard to cervical length and success of labor induction, the current literature demonstrates that cervical length did not accurately predict any outcome, neither a successful vaginal delivery nor a successful cervical ripening [7]. Nonetheless, a differentiation between functional and total cervical length has not been considered.

The type of elastographic method used

As outlined in the introduction, there are to date several elastographic approaches that have been proposed. The most common approach used is the one first proposed by Swiatkowska-Freund and Price in 2011, where strain elastography uses intrinsic motion due to arterial pulsatility and breathing movements [18]. This type of elastography evaluates cervical stiffness in a semi-quantitative color-coded scale. The advantage of this technique is that cervical tissue movements are, in theory, not operator dependent. Intra- and inter-operator evaluation has shown good reliability, even if reliability was evaluated only for the process of strain evaluation, but not for raw dataset acquirement [39]. Nonetheless, as evaluation is only semi-quantitative, significant influences should not be expected.

Another approach, proposed by our study group, targets the quantification of strain obtained by the freehand compression of the cervix through the vaginal probe [40]. The advantage of this method is the quantitative measurement of the strain induced, while the disadvantage is that applied force is operator dependent and that different presetting and setting features could potentially influence the obtained results [41]. Nonetheless, intra- and inter-operator evaluation has shown good reliability after standardization of the force applied, as well as of type of strain used and of the position of the region of interest (ROI) chosen [25, 26].

Interesting is the approach proposed by the Danish group, which proposed to overcome the problem of the exerted force by positioning a reference cup with known Young's modulus on the vaginal probe [42]. Even though the potential advantage of calculating the exact elasticity coefficient of the cervix method should be better standardized as the authors report a good intra-observer intra-class correlation, but a quite poor inter-observer intra-class correlation [20].

To date, only preliminary studies evaluating the practicability of the shear wave technology have been proposed, but no clinical study on the prediction of labor induction success.

The included studies

The recent introduction of cervical elastography has boosted the interest in application of this new tool in the field of obstetrics. Nonetheless, clinical studies are still few. Included in the meta-analysis were only four studies [20–22, 27]. Both in Muscatello et al. and Hwang et al. elastography was performed with the same technique first proposed by Swiatkowska-Freund and Price, using a vaginal probe supported by ElastoScan™ software. Acquisition of a sonoelastogram was performed maintaining the probe steady and using only the movements of the cervix generated by the pulsation of pelvic great vessels and breathing movements of the subject without applying any pressure with the probe.

Two studies (the work of Swiatkowska-Freund and Price and of Pereira et al.) [18, 23] were excluded because it was not possible to retrieve data for the meta-analysis (see Supplemental list 1). One study was excluded as it was not published in the English language [24]. All these studies were performed with the elastographic approach first proposed by Swiatkowska-Freund and Price [18], using ultrasound equipment with ElastoScan™ elastography software (Accuvix XG, Samsung-Medison, Seoul, Korea). The study of Swiatkowska-Freund and Price, including 29 patients, found a significant correlation between elastographic features of the internal os and labor induction success. On the contrary, both the studies of Pereira et al. and Sonnier et al., including respectively a total of 99 and 36 women, found no significant contribution from elastography for prediction of vaginal delivery and induction-to-delivery interval compared to nulliparity and cervical length [23] and for prediction of vaginal delivery within 24 h [24]. Nonetheless, the study by Sonnier et al. could not be evaluated in details as only the abstract was written in English language, and was therefore excluded from the meta-analysis [24]. Furthermore, Pereira et al. found a significant correlation between elastographic score and induction-to-delivery interval in univariate analysis [23]. However, they found this correlation not significant in multivariate analysis but they also stated that as far as there were significant associations between cervical length and the elastographic score, it was possible that studies in larger and more homogeneous populations may demonstrate significant contributions from the elastographic score in the prediction of labor outcome [23].

Discussion on results

For this systematic review and meta-analysis we assessed the efficacy of sonoelastography for prediction of a successful labor induction and vaginal delivery. Pooled diagnostic indexes were assessed for sonoelastography and results compared to the Bishop score (gold standard) and to transvaginal cervical length measurement.

Concerning the diagnostic value of the Bishop score for both labor induction success and vaginal delivery we could consider only the studies of Hwang et al., Hee et al., and Fruscalzo et al. The latter showed a higher positive likelihood ratio compared to the studies of Hwang et al. and Hee et al. [27]. On the contrary, cervical length measurement, again for both endpoints, showed a better prediction ability compared to the Bishop score for all three studies evaluated [20, 21, 27]. Results were also comparable among studies. Interestingly, these results were comparable with results obtained for cervical elastography, where all four included studies were evaluated. Again, results were similar among the considered studies for both endpoints considered.

According to these results we conclude that even though different methods were used for cervical elastography and slightly different endpoints chosen, at least when considering the definition of labor induction success, studies showed a surprising stability of results and superior accuracy than the Bishop score. The Bishop score has a low accuracy in prediction of labor induction success, but it is still the only available system we can use, as sonographic cervical length measurements show inconsistent results. Data show that elastography could be used not only for a more reliable cervical stiffness assessment, but also for a better prediction of labor induction success. This will indeed help in providing better hospital organization and counseling of the patient, as duration can last several days, incurring hospitalization costs and high expectations for the patient. Results should be confirmed in larger clinical trials, and adapted protocols of induction, depending on the cervical ripening status of the patient, could be focused evaluated. Also, a combination of cervical elastography with sonographic cervical length measurement could offer a further tool for a more objective and reliable cervical evaluation than the classic Bishop score now allows.

Limitations

The major bias encountered when performing this meta-analysis was the heterogeneity of the methodology used when performing cervical elastography. Also, the number of studies considered was relative low with a quite limited number of patients included. Furthermore, in two of the excluded studies no advantage of performing cervical elastography was found. These incongruence should be considered and focused evaluated in future studies. However, other biases can limit the interpretation of the results when considering the evaluation of prognostic factors for labor induction success. Among the most relevant that should be mentioned is the heterogeneity of the population considered (parity, gestational age, and presence of PROM), but also the heterogeneity of the protocol used for labor induction (drug type, formulation and dosage applied, use of amniotomy or oxytocin for labor induction) and outcome measures evaluated [27].

Nonetheless, results obtained in this meta-analysis showed a good overlap among studies, particularly when considering cervical elastography and cervical length measurement. Finally, despite these limitations, the intention of this meta-analysis was to evaluate if results are consistent within the studies in order to estimate the potential of cervical elastography for prediction of labor induction success.

Conclusions

Cervix elastography can be considered as a promising and reliable tool for predicting successful labor induction and vaginal delivery in women undergoing a medical induction of labor. Despite the heterogeneity of the methods used, cervical elastography seems to be accurate for the prediction of successful labor induction and vaginal delivery as similar to how ultrasound cervical length measurement was. Furthermore, cervix elastography and cervical length measurement were more reliable in the prediction of a successful labor than the Bishop score.

Financial support: None to declare.

References

[1] ACOG. ACOG Practice Bulletin No. 107: induction of labor. Obstet Gynecol. 2009;114(2 Pt 1):386–97.
[2] Luthy DA, Malmgren JA, Zingheim RW. Cesarean delivery after elective induction in nulliparous women: the physician effect. Am J Obstet Gynecol. 2004;191:1511–5.
[3] Bishop EH. Pelvic scoring for elective induction. Obstet Gynecol. 1964;24:266–8.
[4] Dhall K, Mittal SC, Kumar A. Evaluation of preinduction scoring systems. Aust NZ J Obstet Gynaecol. 1987;27:309–11.
[5] Laughon SK, Zhang J, Troendle J, Sun L, Reddy UM. Using a simplified bishop score to predict vaginal delivery. Obstet Gynecol. 2011;117:805–11.
[6] Teixeira C, Lunet N, Rodrigues T, Barros H. The Bishop Score as a determinant of labour induction success: a systematic review and meta-analysis. Arch Gynecol Obstet. 2012;286:739–53.
[7] Hatfield AS, Sanchez-Ramos L, Kaunitz AM. Sonographic cervical assessment to predict the success of labor induction: a systematic review with metaanalysis. Am J Obstet Gynecol. 2007;197:186–92.
[8] Feltovich H, Hall TJ, Berghella V. Beyond cervical length: emerging technologies for assessing the pregnant cervix. Am J Obstet Gynecol. 2012;207:345–54.
[9] McFarlin BL, Bigelow TA, Laybed Y, O'Brien WD, Oelze ML, Abramowicz JS. Ultrasonic attenuation estimation of the pregnant cervix: a preliminary report. Ultrasound Obstet Gynecol. 2010;36:218–25.
[10] Stein W, Hellmeyer L, Schmidt S, Tekesin I. Intraobserver and interobserver reliability of transvaginal cervical length measurements and quantitative ultrasound tissue characterization of the cervix in the second and third trimester of pregnancy. Ultraschall Med. 2011;32(Suppl 2):E169–74.
[11] Thomas A. Imaging of the cervix using sonoelastography. Ultrasound Obstet Gynecol. 2006;28:356–7.
[12] Hernandez-Andrade E, Hassan SS, Ahn H, Korzeniewski SJ, Yeo L, Chaiworapongsa T, et al. Evaluation of cervical stiffness during pregnancy using semiquantitative ultrasound elastography. Ultrasound Obstet Gynecol. 2013;41:152–61.
[13] Hernandez-Andrade E, Romero R, Korzeniewski SJ, Ahn H, Aurioles-Garibay A, Garcia M, et al. Cervical strain determined by ultrasound elastography and its association with spontaneous preterm delivery. J Perinat Med. 2014;42:159–69.
[14] Kobbing K, Fruscalzo A, Hammer K, Mollers M, Falkenberg M, Kwiecien R, et al. Quantitative elastography of the uterine cervix as a predictor of preterm delivery. J Perinatol. 2014;34:774–80.
[15] Wozniak S, Czuczwar P, Szkodziak P, Milart P, Wozniakowska E, Paszkowski T. Elastography in predicting preterm delivery in asymptomatic, low-risk women: a prospective observational study. BMC Pregnancy Childbirth. 2014;14:238.
[16] Sabiani L, Haumonte JB, Loundou A, Caro AS, Brunet J, Cocallemen JF, et al. Cervical HI-RTE elastography and pregnancy outcome: a prospective study. Eur J Obstet Gynecol Reprod Biol. 2015;186C:80–4.
[17] Feltovich H, Hall TJ. Quantitative imaging of the cervix: setting the bar. Ultrasound Obstet Gynecol. 2013;41:121–8.
[18] Swiatkowska-Freund M, Preis K. Elastography of the uterine cervix: implications for success of induction of labor. Ultrasound Obstet Gynecol. 2011;38:52–6.
[19] Fruscalzo A, Londero AP, Frohlich C, Meyer-Wittkopf M, Schmitz R. Quantitative elastography of the cervix for predicting labor induction success. Ultraschall Med. 2015;36:65–73.
[20] Hee L, Rasmussen CK, Schlutter JM, Sandager P, Uldbjerg N. Quantitative sonoelastography of the uterine cervix prior to induction of labor as a predictor of cervical dilation time. Acta Obstet Gynecol Scand. 2014;93:684–90.
[21] Hwang HS, Sohn IS, Kwon HS. Imaging analysis of cervical elastography for prediction of successful induction of labor at term. J Ultrasound Med. 2013;32:937–46.
[22] Muscatello A, Di Nicola M, Accurti V, Mastrocola N, Franchi V, Colagrande I, et al. Sonoelastography as method for preliminary evaluation of uterine cervix to predict success of induction of labor. Fetal Diagn Ther. 2014;35:57–61.
[23] Pereira S, Frick AP, Poon LC, Zamprakou A, Nicolaides KH. Successful induction of labor: prediction by preinduction cervical length, angle of progression and cervical elastography. Ultrasound Obstet Gynecol. 2014;44:468–75.
[24] Sonnier L, Bouhanna P, Arnou C, Rozenberg P. Elastography of cervix to predict delay from induction to delivery. Gynecol Obstet Fertil. 2014;42:827–31.
[25] Fruscalzo A, Schmitz R, Klockenbusch W, Steinhard J. Reliability of cervix elastography in late first and second trimester of pregnancy. Ultraschall Med. 2012;33:1–7.
[26] Fruscalzo A, Steinhard J, Londero AP, Fröhlich C, Bijnens B, Klockenbusch W, et al. Reliability of quantitative elastography of the uterine cervix in at-term pregnancies. J Perinat Med. 2013;41:421–7.
[27] Fruscalzo A, Londero AP, Frohlich C, Mollmann U, Schmitz R. Quantitative elastography for cervical stiffness assessment during pregnancy. Biomed Res Int. 2014:826535. doi: 10.1155/2014/826535.
[28] Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2009. p. 1–27.
[29] Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol. 2006;6:9.
[30] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Br Med J. 2003;327:557–60.
[31] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Br Med J. 1997;315:629–34.

[32] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994;50:1088–101.
[33] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. J Am Med Assoc. 2000;283:2008–12.
[34] Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6:e1000097.
[35] Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Br Med J. 2011;343:d4002.
[36] Lange AP, Secher NJ, Westergaard JG, Skovgard I. Prelabor evaluation of inducibility. Obstet Gynecol. 1982;60:137–47.
[37] Hoesli IM, Strutas D, Tercanli S, Holzgreve W. Charts for cervical length in singleton pregnancy. Int J Gynaecol Obstet. 2003;82:161–5.
[38] Londero AP, Bertozzi S, Fruscalzo A, Driul L, Marchesoni D. Ultrasonographic assessment of cervix size and its correlation with female characteristics, pregnancy, BMI, and other anthropometric features. Arch Gynecol Obstet. 2011;283:545–50.
[39] Fruscalzo A, Schmitz R. Reply. Ultrasound Obstet Gynecol. 2013;41:712–4.
[40] Fruscalzo A, Schmitz R. Quantitative cervical elastography in pregnancy. Ultrasound Obstet Gynecol. 2012;40:612.
[41] Fruscalzo A, Londero AP, Schmitz R. Quantitative cervical elastography during pregnancy: influence of setting features on strain calculation. J Med Ultrason. 2015; DOI 10.1007/s10396-015-0619-3.
[42] Hee L, Sandager P, Petersen O, Uldbjerg N. Quantitative sonoelastography of the uterine cervix by interposition of a synthetic reference material. Acta Obstet Gynecol Scand. 2013;92:1244–9.

The authors stated that there are no conflicts of interest regarding the publication of this article.

Supplemental Material: The online version of this article (DOI: 10.1515/jpm-2015-0035) offers supplementary material, available to authorized users.