HHS Public Access

Author manuscript

Am J Obstet Gynecol. Author manuscript; available in PMC 2016 February 01.

Published in final edited form as:

Am J Obstet Gynecol. 2015 February ; 212(2): 145–156. doi:10.1016/j.ajog.2014.07.003.

Precocious Cervical Ripening as a Screening Target to Predict Spontaneous Preterm Delivery among Asymptomatic Singleton Pregnancies: A Systematic Review

Qing Li, M.D., Dr.P.H.[1], **Mathew Reeves, B.V.Sc., Ph.D.**[1], **John Owen, M.D.**[2], and **Louis Keith, M.D., Ph.D.**[3]

[1]Department of Epidemiology & Biostatistics, Michigan State University

[2]Department of Obstetrics and Gynecology, University of Alabama at Birmingham

[3]Department of Obstetrics and Gynecology, Northwestern University

Abstract

Routine second-trimester transvaginal ultrasonographic (TVU) screening for short cervical length (CL) predicts spontaneous preterm delivery (SPTD), albeit with limited sensitivity (35–40%) and moderate positive likelihood ratio (LR+: 4–6). However, CL describes one of multidimensional changes associated with precocious cervical ripening (PCCR), which also include cervical softening, cervical funneling (CF), and dilation. PCCR, a precursor and a strong predictor for SPTD, was proposed as a potential screening target. We hypothesized that screening for composite measures of PCCR (e.g. CL, CF, cervical consistency, and dilation) using either digital exam (DE) or TVU would improve prediction of SPTD compared to screening for short CL alone. We searched PubMed and EMBASE electronic databases for observational cohort studies to evaluate cervical screening in asymptomatic obstetric populations. Multidimensional composite cervical measures were assessed in 10 datasets (n=22,050 pregnancies) and 12 publications. Appreciable heterogeneity in cervical measurements, data quality, and outcomes across studies prevented quantitative meta-analysis. Only one study reported intra- and inter-observer reliability of cervical measurements. The prevalence of CF ranged from 0.7% to 9.1%. Five studies compared composite measures of PCCR (i.e., CL and CF) to short CL alone, and consistently reported improved screening performance. Among three TVU studies, gains in sensitivity ranged from 5% to 27%, and increases in LR+ ranged from 3 to 16. Our findings suggest composite measures of PCCR might serve as valuable screening targets. High-quality interdisciplinary studies integrating epidemiologic approaches are needed to test this hypothesis and accelerate the translation of advances in cervical pathophysioloy into effective preventive interventions.

Correspondence and contact for reprint request: Qing Li, MD, DrPH, NICHD T-32 Perinatal Epidemiology program, Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI 48824, Phone: +1-205-427-7556, qli@epi.msu.edu; youliqing08@gmail.com.

Disclosure Statement: The authors report no conflict of interest.

Keywords

Cervical Screening; Preterm Delivery; Precocious Cervical Ripening; Precursor; Epidemiologic Approaches; Translational Research

INTRODUCTION

Spontaneous preterm delivery (SPTD) is an unsolved public health problem of global proportions,[1] requiring more effective prevention strategies.[2, 3] Timely prevention of SPTD commences with early identification of a modifiable target by means of effective screening programs.[4, 5] In general, short cervical length (CL) is the screening target in routine second-trimester transvaginal ultrasonography (TVU, Figure 1, left and middle), which represents a simple, safe, and reproducible technical advance,[6, 7] compared to digital examination (DE). While TVU has distinct advantages, DE doesn't require sonographic training or equipment, and may therefore be more suitable for resource-limited settings. Whereas the false-positive rates of TVU and DE screening are similar,[8] the limited sensitivity (35–40%) of TVU screening for shortened CL is marginally better than that of DE (25–30%).[2, 8] Considering short CL, defined at less than 25 mm at <20 weeks, a UK review reported a moderate LR+ of 6.29 to predict SPTD before 34 weeks,[9] and a Canadian review reported a LR+ of 4.31 to predict SPTD before 35 weeks.[10] Moderate LR+s of short cervices assessed between 20–24 weeks were consistently reported,[11, 12] including a value of 2.86 for asymptomatic high-risk women with a SPTD history.[10]

If the obstetric community had an effective and efficient means of screening for SPTD, this approach could be expanded to routine use in all pregnant women. Previously, there was a lack of evidence for the value of early intervention. However, a recent meta-analysis demonstrated that vaginal progesterone administration to asymptomatic women with a sonographic short cervix not only reduced the risk of SPTD but also led to a 43% reduction in neonatal morbidity and mortality.[13] Although universal screening for short cervices followed by progesterone treatment is cost-effective,[14–16] large (400 or 588) numbers of mothers must be screened to prevent one SPTD.[17] Clearly, more efficient screening strategies are needed. According to The American Congress of Obstetricians and Gynecologists Practice Bulletins (#101 in 2009 and #130 in 2012), the maternal cervix "should be examined as clinically appropriate when technically feasible;"[18(p452)] and universal cervical screening of pregnant women without a prior preterm birth may be considered despite "limited or inconsistent scientific evidence (Level B recommendation)."[7(p970)] Furthermore, evidence-based research is required for greater quality assurance.[7] Although the U.S. Preventive Services Task Force acknowledges the importance of predicting preterm delivery through screening, it has not recommended any screening targets.[19–21]

Theory and Reasoning for Prediction

For years, multidimensional cervical features were used to predict the early onset of labor or SPTD. In 1964, the Bishop scoring system (cervical dilation, effacement, consistency, and position as assessed by DE), was correlated with the time to the onset of labor.[22] In 1965,

Wood et al. first reported that an internal cervical os dilated to one finger breadth and an effaced cervix predicted SPTD.[23] Papiernik and colleagues reported a decline in SPTD prior to 32 weeks (1.7% in 1971–1974 *vs* 0.8% in 1979–1982) in the French city of Haguenau after implementation of uncontrolled and multilevel interventions.[24, 25]Prominent among targets of this population-based risk assessment and screening system[25] were both shortened cervices and patency of the internal os.[26] These precocious signs of cervical ripening can be recognized during a vaginal examination several weeks before the onset of SPTD and may be useful in predicting it.[27] Despite the presentation of this French screening experience at conferences[28, 29] and in a March of Dimes monograph[24] aimed to the U.S. medical audience,[30] this paper published in 1986[27] hasn't been cited widely in three past decades (78 citations on Web of Knowledge and 102 on Google Scholar in March 2014), and deserves a new look. Furthermore, we acknowledge again recent progress in available effective treatments, such as vaginal progesterone,[7] which is an essential criterion required to support screening.[4, 5, 31]

Identifying effective screening targets for SPTD relies on an understanding of its natural history and pathophysiology; in the latter circumstance, our understanding is lacking. Because precocious cervical ripening (PCCR) is an important precursor state in the SPTD pathway and a strong predictor for it, PCCR is a potential target for screening. Precursors are pathologic states that have a high probability of progressing to disease after a latent stage.[32] Accordingly, ascertainment of properly defined precursors can increase the effectiveness of screening and prevention.[32] As a recognizable stage in parturition,[33–35] the term PCCR was initially coined by Papiernik and colleagues in 1986.[27] PCCR describes multidimensional cervical changes including softening, shortening, funneling, and dilation of the internal os. These changes, visible using ultrasound,[36] progress from T-, to Y-, V- or U-shape funnels (Figure 1, right) prior to the onset of SPTD.[27, 37] Cervical pathophysiology has been further investigated through molecular and cellular approaches.[38–40] Romero and colleagues described cervical ripening as a general feature of the "premature parturition syndrome."[41, 42] In 2011, routine recording of cervical ripening was recommended by the Global Alliance to Prevent Prematurity and Stillbirth.[43] In 2012, Caritis and Simhan proposed that the term PCCR was more appropriate and less confusing than either "cervical incompetence"[6, 44, 45] or "cervical insufficiency,"[46] both being ill-defined biologically.[47] In this review, we use the term PCCR and operationalize it as at least two measurable cervical dimensions.

It is logical to ask how well the performance of PCCR has been evaluated to date in predicting SPTD. The effectiveness of a screening program depends on the interrelations between: 1) the performance, timing and frequency of screening procedures; 2) the efficacy of timely interventions; and 3) the risk profile of target populations.[35] We chose to investigate both reviews as well as individual studies; but we confine our comments regarding reviews to the introduction. Reviews by Owen and colleagues[44] and Honest and colleagues[9, 12] grouped only observational studies; other reviews mixed clinical trials and observational studies together.[48–50] Despite providing useful insights concerning diverse populations, designs and analytical methods, prior reviews[9–12, 17, 44, 48–52] failed to consider PCCR with most investigators focusing entirely on CL as measured by TVU.[17, 44, 49–51]

Reports from five investigative teams[9–12, 48, 52, 53] over the past 15 years did not cite Papiernik et al.'s screening paper from 1986[27] but considered some of the hypotheses which form the basis for the present analysis. The first was Leitich and colleagues from Austria who concluded dilatation of the internal cervical os to be among the most effective markers for preterm delivery.[52] The second was Honest and colleagues from the UK who published three reviews[9, 11, 12] and reported that 1) the larger the funnel (e.g., dilatation of internal os >5 mm), the more accurately the prediction of SPTD; and 2) CL and cervical funneling (CF), used alone or in combination, appeared useful for SPTD prediction, but no data were highlighted. The third team was Crane and Hutchens from Canada, who included CF in their tables but did not summarize its predictive performance.[10] The fourth team was Reiter and colleagues from Denmark, who published the only review that chose to target "premature cervical ripening" and reported unclear methods for the estimation and the insufficient evidence for routine screening;[48] however, they neglected to justify this target and to include CF from studies, such as the one from Iams and colleagues.[8] Finally, Barros-Silva and colleagues from Portugal reported inconsistent findings in comparing combined screening targets to short CL alone in three studies, and recommended combining CL "with other markers (sonographic, biochemical and/or clinical) that reflect the multiplicity of mechanisms involved in the pathogenesis of SPTD".[53(5)]

Objective

We hypothesized that comprehensive assessment for multidimensional PCCR (e.g., CL, CF, cervical consistency, and cervical dilation in combination) is more effective (e.g., improved sensitivity and LR+) than screening for short CL alone using either TVU or DE.[54] DE is suitable for resource-limited settings, serves as a historical comparison, and is included. The primary outcome measure was SPTD at specified gestational weeks. In this systematic review, we aimed to identify, appraise, select and synthesize all high-quality research evidence. We assessed elements of study methodology and variations in cervical assessment, risk profiles of participants, and healthcare contexts. Further, we identified research gaps and suggest future research to improve the performance of cervical screening and its cross-cultural applicability.

METHODS FOR REVIEW

Selection criteria and data sources

A comprehensive literature search was conducted to identify articles published in English language journals from 1980 to March 2014. We included high-quality studies, which evaluated multidimensional aspects of PCCR to predict singleton SPTDs in observational cohort studies of unselected obstetric populations. We excluded studies assessing *only one* cervical dimension, i.e., CF[55] or CL.[56–58] This systematic review included only published manuscripts and was therefore exempt from Institutional Review Board (IRB) review.

Using the key words *Cervi*, Preterm, Prematurity*, and *Predict*, singleton**, we searched PubMed and EMBASE electronic databases in March, 2014. We identified 538 reports and two from other sources as depicted in a flow diagram[59] (Figure 2) from the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA);[60] 397 citations and

abstracts yielded 160 potentially relevant articles for full-text review. We excluded 119 papers considering only one cervical dimension and 29 papers including multiple cervical measures. Reference lists were manually searched but failed to reveal other studies. Twelve reports describing ten datasets comprising 22,050 pregnancies met the inclusion criteria.[27, 61–65]

Screening is the identification of asymptomatic disease or risk factors.[66] Assessing high-risk (e.g., mothers with a history of SPTD) or symptomatic mothers is not an appropriate design for comparative screening evaluation and will limit the generalizability of findings to other populations. We have been prudent to the inclusion and exclusion criteria so that we can identify all high-quality evidence of screening effectiveness. Twenty-five of the 29 excluded studies (citations are available upon request) assessed cervices of high-risk (n=4) or symptomatic (n=21) mothers; one excluded women with the history of SPTD;[67] Among three studies with asymptomatic participants, two did not evaluate predictive performances of cervical assessment,[68, 69] and one evaluated cervical measures other than cervical length.[70]

Measurements and evaluation process

The two essential characteristics of a screening test are its reliability and validity.[66] Therefore, when available we abstracted data on reliability, sensitivity, specificity, positive and negative predictive values (PPV, NPV), positive and negative likelihood ratios (LR+, LR-), and receiver operating characteristic (ROC) curves.[71] However, due to the variation in reporting across studies, we only focused on reporting sensitivity and LRs. We calculated LRs based on the following formulas.[66]

$$LR+=\frac{\text{sensitivity}}{1-\text{specificity}}\text{LR}-=\frac{1-\text{sensitivity}}{\text{specificity}}$$

To assess study quality, we used the recommendations from the Standards for Reporting of Diagnostic Accuracy (STARD) Steering Group[72] and a review of screening tests[73] to document the following study criteria: sample characteristics (consecutive sample), study design (prospective or retrospective cohort), cervical assessment (evaluator background, sonographers, obstetricians, etc.), blinding, reliability, outcomes (definition of SPTD), and quantitative analysis (screening performance, and statistical association between cervical measures and SPTD).

The use of multidimensional cervical changes in the inclusion criteria allowed us to evaluate the performance of individual PCCR components. First, within a given population, we calculated the difference in sensitivity and/or LRs of composite measures of PCCR compared to short CL alone. Second, we assessed the consistency of within-study comparisons across studies.[74] However, a quantitative meta-analysis could not be performed due to 1) the lack of sensitivity and LRs data or data required to reproduce the 2×2 contingency tables and 2) the appreciable heterogeneity in design, screening, definitions of CF, and variable cutoff values.

RESULTS

General time and geographic trends

The 12 studies reviewed were published at a rate of approximately 3–4 per decade in the past thirty years (Table 1). They were conducted in Europe (5 publications, 41%), North America (3 publications, 25%), Asia (2 publications, 17%), and South America (2 publications, 17%). Most were from high-income countries (i.e., France, Sweden, Japan, Finland, the U.S., and the United Kingdom) or high-income regions (i.e., Hong Kong), but two originated in middle-income countries (i.e., Brazil and Colombia).

Study quality

All studies used prospective cohorts with consecutive cases (Table 2). The patients and providers in two studies and providers in five studies were blinded to cervical measures.[8, 62, 63, 75–78] Intra/inter observer reliability in studies of ultrasound measures was poorly described; despite three studies claiming to have used rigorous quality control processes,[77, 79, 80] only one reported the intra- and inter-observer reliability of the cervical consistency index.[80] Older studies reported associations (e.g., adjusted relative risk) between cervical measurements and the risk of SPTD.[27, 61, 64] However, two of these did not include measures of sensitivity, specificity, or LRs,[27, 64] and three others reported only a single criterion (i.e., sensitivity or PPV).[76, 80, 81]

Study populations and definitions of SPTD

The low incidence rates of SPTD (e.g., <9% before 37 weeks, <5% before 34 or 35 weeks, and <1% before 33 weeks; Table 3) reflected generally low-risk populations. Mothers were recruited from multi-center studies[8, 75, 77, 81] or hospitals.[76, 80] Only hospitals from Finland, France, Hong Kong, Sweden, and the UK integrated cervical assessment into routine prenatal service an institutionalized preventive intervention to predict SPTD.[8, 27, 61–65, 75–77, 80, 81]

Cervical assessment specifics

The timing, mode, and training of evaluators varied across studies (Table 2). Four studies screened only once during pregnancy,[62–64, 76] two screened twice,[8, 77] and one screened three times.[65] In studies conducted in the 1980s, obstetricians or midwives performed DE only. Three American studies evaluated DE by obstetricians and nurses[75] or by nurses and standard examiners[8, 77] who had at least 5 years of experience in cervical examinations and were designated as the “standard” to which all cervical examiners were compared.[77] However, since the 1990s TVU, as performed by sonographers, obstetricians and midwives, has been common. No harmful effect of either method was reported.

Measurements and screening performance

The evaluation of multidimensional cervical screening (Table 3) began with the French study on PCCR,[27] which was followed by three others using the 5-component Bishop score or modifications thereof.[8, 65, 77] Overall, insufficient standardization exists in terms of methods and definitions to compare performance criteria across studies. For example, the

prevalence of CF ranged from 0.7% to 9.1%. Using data from the Maternal Fetal Medicine Unit (MFMU) Network, sensitivity was improved (e.g., 49.4% vs 37.3% for CL≤ 25 mm) at the second cervical assessment (26–28 weeks) compared to the first (22–24 weeks).[8, 77] The change of performance over three assessments[65] was not reported.

The focus in cervical evaluation within a single screening section changed from their associations with SPTD to their predictive performance of SPTD. Most studies reported significant associations between cervical dilation and later CF and SPTD (Table 3). Specifically, Papiernik and colleagues[27] and Mortensen and colleagues[65] reported cervical dilation as being the strongest predictor for SPTD compared to other assessable dimensions. Bouyer and colleagues[61] reported that cervical dilation had a persistently higher relative risk among parous women compared to nulliparous women. In contrast, To and colleagues noted that the significant association between CF and SPTD in univariate analysis became null after adjustment for CL.[64] As shown in Table 3, three studies reported larger LR+s for cervical dilation or CF compared to that of short CL alone (range of difference: 0.2–9.7),[8, 63, 75] and two studies reported smaller LR+s (0.2–4.6).[8, 62] Despite null findings from studies performed in the MFMU Network,[8, 77] one reported better predictive performance of screening in parous mothers,[61] whereas the other suggested improved performance in nulliparous mothers.[81]

Using both CL and CF (i.e., either one being abnormal, both abnormal, or a combined score), only five of 12 studies compared composite measures of PCCR to that of a short CL alone within a single screening section, and reported gains in sensitivity (from 4% to 27%) and LR+ (1 for DE and comparing DE to TVU, and from 3.4 to 16 for TVU) (Table 4). Hartmann and colleagues[75] and Newman and colleagues[77] used a cervical score (CL minus dilation in centimeters) based on DE (i.e., sensitivity: 20% *vs* 13%, 36% *vs* 32%; LR+: 2.9 *vs* 1.9, 7.4 *vs* 6.4). Using TVU, Taipale and Hiilesmaa[63] used a measure of either a short CL or cervical dilation (sensitivity: 29% *vs* 19%; LR+: 9.7 *vs* 6.3) to predict SPTD. Leung and colleagues,[62] combining CL and CF, improved LR+ (26 *vs* 9.8) but reduced sensitivity (26% *vs* 37%), whereas using either indicator improved sensitivity (42% *vs* 37%), but reduced LR + (4.7 *vs* 9.8). De Carvalho and colleagues[76] combined CF with short CL and improved sensitivity (34% *vs* 7%).

We found other targets of cervical measures in addition to CL and CF and the consideration of parity. One study investigated cervical consistency, the ratio of the anteroposterior cervical diameter measured before and after application of pressure on the cervix using the transvaginal probe, multiplied by 100.[80] Another study compared the performance of summary scores through stratifying by parity.[61] Two studies[75, 77] applied the formula of a cervical score validated among twin gestations[82–84] to singleton gestations.

COMMENT

This review systematically reviewed the available evidence from three decades of comprehensive cervical screening for PCCR in large observational cohorts of asymptomatic singleton gestations to predict SPTD. Only studies from Finland, France, Hong Kong, Sweden, and the UK integrated cervical assessment into routine prenatal care as a standard.

Most of the studies reviewed had insufficient standardization and varied by methodological quality, outcomes (i.e. early or all SPTD), and cervical assessment (i.e. dimensional measurements, and timing and frequency of examination). Shorter CL and CF had high specificity, but low sensitivity to predict SPTD. All five studies that used composite measures consistently showed improved predictive performance compared to those which used CL alone. Taken together, our findings indicate that composite measures of PCCR could represent valuable targets of future research to optimize the prediction of SPTD.

Existing studies are limited in their scope, and do not include reports from low-income countries.[85] Variable screening performance could be explained by *measurement issues* including methods, evaluators, facilities, and global populations. For example, regional- and racial-specific cervical assessments (e.g., gestational end points for racial groups across continents) were reported.[62, 86–88] Blinded cervical evaluation was not widely used. Standardization of cervical assessment[89] is necessary to improve the validity of cervical screening. Despite the difficulty in obtaining reliable CF measures,[58, 90] such details need further investigation and quantification to improve their reliability. Only one study reported reliability.[80] Future high-quality studies should assess intra-rater and inter-rater reliability of cervical measurements and evaluate reasonable cervical predictors to generate empirical evidence.

We also determined that the lack of the underlying theory and the logic behind prediction might explain the varied performances across screening tests. As stated in our introduction, the features of an ideal screening test start with a detectable target based on pathophysiology. PCCR is an important but under-researched concept; only five groups[9–12, 46, 50, 51] considered some of the hypotheses which form the basis for the present analysis. Using an operational definition of PCCR and available empirical measurements, we found evidence that composite measures of PCCR improved predictive performance compared to CL alone.[62, 63, 75–77] and may thus serve as valuable potential screening targets.

Our review also discovered *analytical issues* that might explain observed variances in screening performance. Few studies used LR±[62] and ROC curves.[77] To optimize analytical processes, we propose a 2-step multivariable prediction model.[91, 92] First, identify and model measurements to generate a formula. Second, quantify scores and cut-off points based on the best performance,[66] evaluating the effect of adding additional characteristics to a short cervix.[93] It is necessary to consider profiles of target populations. Originally, To and colleagues[64] reported an insignificant association between CF and SPTD, but their data suggested an additive value of CF depending on CL (Figure 3). The correlation between CF and CL and the potential multicollinearity[94] are worthy of consideration.

Despite strengths, including “within-study comparisons” and aligning the pathophysiology, definition and function of PCCR as a precursor to optimize the prediction of SPTD, this review is subject to several limitations. First, related reports may not have been identified if they did not include key words in abstracts or titles. Second, we could not access original data to test whether observed differences in screening performance were significantly different. In this regard, we could neither perform individual patient data meta-analyses, nor

synthesize cut-off points of measures across studies (e.g., definitions of SPTD, features of CF, percentile values or Z-scores for CL).[95] Third, the generalizability of the findings from studies using convenience samples of women is limited. Finally, repeat cervical screening,[96–98] screening coupled with interventions,[54] the concurrent use of other predictors (e.g., collagen structure, elasticity,[99] consistency,[80] fetal fibronectin,[100] amniotic fluid sludge,[101, 102] maternal position,[103, 104] and multiple gestations[103]) as well as cost effectiveness deserve future attention.

CONCLUSIONS

Further research on comprehensive cervical assessment for multidimensional PCCR to predict singleton SPTD is justified for several reasons:

First, the recent increase in global attention to SPTD invites the development of clinical innovations with potential for primary and secondary prevention. Although the concept of PCCR is grounded in clinical tradition, its application and empiric measurement in screening tests require innovative and epidemiologic approaches to generate contemporary evidence.

Second, important gaps must be addressed in the preparation of an optimal, evidence-based protocols and high-quality comparative studies on screening for PCCR. Key unknowns in this daunting task include the lack of data on analytical approaches to incorporate cervical measurements, timing, frequency, and reliability of screening. In this regard, secondary data analyses also can be valuable. Benefits and harms should be assessed (e.g., under- or over-prediction).[5] Standardized research is required to improve the conceptualization, measurement, and validation of comprehensive cervical screening for PCCR.

Third, the U.S. has much to contribute to global initiatives to predict and prevent SPTD, as it contributes half of all SPTD cases among high-income countries.[85] In 1986, Papiernik noticed the lack of routine cervical assessment in both the U.S. and Great Britain compared to continental Europe;[27] whereas the United Kingdom adopted it later in 1997,[64] the U.S. has not and high-quality studies[8, 75, 77] included in this review were designed 20 years ago. Today's pressing needs include the national approaches to advance research and practice using high-quality design and new data. In this regard, cervical data are being collected from 10,000 nulliparous American women from eight sites in a research network.[105]

Finally, high-quality studies using an interdisciplinary approach including epidemiology are needed to test the hypothesis of PCCR as a target and accelerate the translation of advances in pathophysioloy into effective preventive interventions. This journal previously has called for accelerating efforts in collaborative and translational research.[106] By synthesizing knowledge across disciplines (e.g., cervical pathophysiology, clinical epidemiology, and maternal fetal medicine), epidemiology can play a central role and provide methods and tools to enhance translational research[107] and facilitate evidence-based practice.[106, 108] Precursors,[32] predisease,[109] and "predictor of poor health"[5] can advance preventive interventions, a successful example being cervical intraepithelial neoplasia grade 3 detection for cervical cancer prevention.[32]

Acknowledgments

Qing Li's data abstraction and preparation of this review were supported by an Institutional T32 grant in Perinatal Epidemiology from the National Institute of Child Health and Human Development at the National Institutes of Health (Grant no. HD046377/HD/NICHD NIH HHS) awarded to the Principal Investigator Nigel Paneth, Michigan State University. The authors gratefully acknowledge comments from three reviewers invited by the American Journal of Obstetrics & Gynecology, Drs. Nigel Paneth, Claudia Holzman, Zhiying You, and the Pregnancy Outcomes and Community Health (POUCH) team from Michigan State University, Jay Iams from Ohio State University, and Birgit Arabin from Center for Mother and Child of the Philipps University Marburg, and Clara Angela Foundation on an early draft of this manuscript. Preliminary study findings were presented at the Society for Epidemiologic Research and the Society for Pediatric and Perinatal Epidemiologic Research in Boston in June 2013.

Abbreviations

CF	cervical funneling
CL	cervical length
DE	digital examination
MFMU	Maternal Fetal Medicine Unit
PCCR	precocious cervical ripening
PPV	positive predictive values
LR+	positive likelihood ratio
SPTD	spontaneous preterm delivery
ROC curves	receiver operating characteristic curves
TVU	transvaginal ultrasonography

References

1. Holzman C, Paneth N. Preterm birth: from prediction to prevention. Am J Public Health. 1998; 88:183–184. [PubMed: 9491003]
2. Iams JD, Romero R, Culhane JF, Goldenberg RL. Primary, secondary, and tertiary interventions to reduce the morbidity and mortality of preterm birth. Lancet. 2008; 371:164–175. [PubMed: 18191687]
3. Howson, C.; Kinney, M.; Lawn, JE., editors. Born Too Soon: The Global Action Report on Preterm Birth. Geneva: World Health Organization; 2012. March of Dimes; the Partnership for Maternal Newborn and Child Health; Save the Children; World Health Organization. Accessed Available from: http://www.who.int/pmnch/media/news/2012/201204_borntoosoon-report.pdf [Retrieved September 7, 2012]
4. Holland, WW.; Stewart, S. Screening in Disease Prevention: What Works?. Oxford: Radcliffe Publishing Ltd in association with The Nuffield Trust and the European Observatory on Health Systems and Policies;
5. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: reflections from 4 current and former members of the U.S. Preventive services task force. Epidemiol Rev. 2011; 33:20–35. [PubMed: 21666224]
6. Owen J, Iams JD. National Institute of Child H, Human Development Maternal-Fetal Medicine Units N. What we have learned about cervical ultrasound. Semin Perinatol. 2003; 27:194–203. [PubMed: 12889586]
7. Acog Committee on Practice Bulletins. ACOG Practice Bulletin No. 130: Prediction and prevention of preterm birth. Obstet Gynecol. 2012; 120:964–973. [PubMed: 22996126]

8. Iams JD, Goldenberg RL, Meis PJ, et al. The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. N Engl J Med. 1996; 334:567–572. [PubMed: 8569824]
9. Honest H, Bachmann LM, Coomarasamy A, Gupta JK, Kleijnen J, Khan KS. Accuracy of cervical transvaginal sonography in predicting preterm birth: a systematic review. Ultrasound Obstet Gynecol. 2003; 22:305–322. [PubMed: 12942506]
10. Crane JM, Hutchens D. Transvaginal sonographic measurement of cervical length to predict preterm birth in asymptomatic women at increased risk: a systematic review. Ultrasound Obstet Gynecol. 2008; 31:579–587. [PubMed: 18412093]
11. Honest H, Forbes CA, Duree KH, et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. Health Technol Assess. 2009; 13:1–627. [PubMed: 19796569]
12. Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: no good test for predicting a spontaneous preterm birth. Curr Opin Obstet Gynecol. 2012; 24:422–433. [PubMed: 23099810]
13. Romero R, Nicolaides K, Conde-Agudelo A, et al. Vaginal progesterone in women with an asymptomatic sonographic short cervix in the midtrimester decreases preterm delivery and neonatal morbidity: a systematic review and metaanalysis of individual patient data. Am J Obstet Gynecol. 2012; 206:124 e1–124 e19. [PubMed: 22284156]
14. Cahill AG, Odibo AO, Caughey AB, et al. Universal cervical length screening and treatment with vaginal progesterone to prevent preterm birth: a decision and economic analysis. Am J Obstet Gynecol. 2010; 202:548 e1–548 e8. [PubMed: 20079888]
15. Werner EF, Han CS, Pettker CM, et al. Universal cervical-length screening to prevent preterm birth: a cost-effectiveness analysis. Ultrasound Obstet Gynecol. 2011; 38:32–37. [PubMed: 21157771]
16. Odibo AO, Stamilio DM, Macones GA, Polsky D. 17alpha-hydroxyprogesterone caproate for the prevention of preterm delivery: A cost-effectiveness analysis. Obstet Gynecol. 2006; 108:492–499. [PubMed: 16946206]
17. Parry S, Simhan H, Elovitz M, Iams J. Universal maternal cervical length screening during the second trimester: pros and cons of a strategy to identify women at risk of spontaneous preterm delivery. Am J Obstet Gynecol. 2012; 207:101–106. [PubMed: 22633270]
18. American College of Obstetricians Gynecologists. ACOG Practice Bulletin No. 101: Ultrasonography in pregnancy. Obstet Gynecol. 2009; 113:451–461. [PubMed: 19155920]
19. The U.S. Preventive Services Task Force. Screening for Bacterial Vaginosis in Pregnancy to Prevent Preterm Delivery: The U.S. Preventive Services Task Force, February 2008. Accessed Available from: http://www.uspreventiveservicestaskforce.org/uspstf/uspsbvag.htm#clinical.
20. The U.S. Preventive Services Task Force. Screening for Ultrasonography in Pregnancy: The U.S. Preventive Services Task Force, 1996. Accessed Available from: http://www.uspreventiveservicestaskforce.org/uspstf/uspsuspg.htm.
21. U.S. Preventive Services Task Force. Screening for bacterial vaginosis in pregnancy to prevent preterm delivery: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2008; 148:214–219. [PubMed: 18252683]
22. Bishop E. Pelvic scoring for elective induction. Obstet Gynecol. 1964; 24:266–268. [PubMed: 14199536]
23. Wood C, Bannerman RH, Booth RT, Pinkerton JH. The Prediction of Premature Labor by Observation of the Cervix and External Tocography. Am J Obstet Gynecol. 1965; 91:396–402. [PubMed: 14259927]
24. Papiernik, E.; Keith, L.; Bouyer, J.; Dreyfus, J.; Lazar, P. Effective Prevention of Preterm Birth: The French Experience Measured at Haguenau. White Plains, NY: March of Dimes Birth Defects Foundation;
25. Papiernik E, Bouyer J, Dreyfus J, et al. Prevention of preterm births: a perinatal study in Haguenau, France. Pediatrics. 1985; 76:154–158. [PubMed: 4022686]
26. Papiernik E, Kaminski M. Multifactorial study of the risk of prematurity at 32 weeks of gestation. I. A study of the frequency of 30 predictive characteristics. J Perinat Med. 1974; 2:30–36. [PubMed: 4468289]

27. Papiernik E, Bouyer J, Collin D, Winisdoerffer G, Dreyfus J. Precocious cervical ripening and preterm labor. Obstet Gynecol. 1986; 67:238–242. [PubMed: 3945433]
28. Papiernik, E.; Spira, N.; Bréart, G., et al. Pre´vention de la Naissnace Pre´mature´e : Nouveaux objectifs et nouvelles pratiques des soins pre´nataux (Prevention of preterm birth: New goals and new practices in prenatal care): colloque INSERM, Evian, 19–22 mai 1985. Paris, France: INSERM;
29. Papiernik E, Charlemain C, Goffinet F, Paul G, Keith L. Vaginal bacterial colonization, uterine cervix and preterm births. Prenatal and Neonatal Medicine. 1998; 3:98–102.
30. Alexander GR, Weiss J, Hulsey TC, Papiernik E. Preterm birth prevention: an evaluation of programs in the United States. Birth. 1991; 18:160–169. [PubMed: 1930445]
31. Wald, N.; Law, M. Medical screening. In: Warrell, DA.; Cox, TM.; Firth, JD., editors. Oxford Textbook of Medicine. Vol. 1. New York, NY: Oxford University Press; 2010.
32. Wacholder S. Precursors in cancer epidemiology: Aligning definition and function. Cancer Epidemiology, Biomarkers & Prevention. 2013; 22:521–527.
33. Challis JRG. Mechanism of parturition and preterm labor. Obstet Gynecol Surv. 2000; 55:650–660. [PubMed: 11023206]
34. Preterm Birth: Causes, Consequences, and Prevention. Washington DC: National Academy of Sciences; Institute of Medicine Committee on Understanding Premature Birth Assuring Healthy Outcomes.
35. Morrison, AS. Screening. In: Rothman, KJ.; Greenland, S., editors. Modern Epidemiology. Philadelphia, PA: Lippincott-Raven Publishers; 1998.
36. Berghella, V.; Bega, G. Ultrasound evaluation of the cervix. In: Callen, PW., editor. Ultrasonography in Obstetrics and Gynecology. Philadelphia, PA: Saunders Elsevier; 2008.
37. Zilianti M, Azuaga A, Calderon F, Pages G, Mendoza G. Monitoring the effacement of the uterine cervix by transperineal sonography: a new perspective. J Ultrasound Med. 1995; 14:719–724. [PubMed: 8544236]
38. Word RA, Li XH, Hnat M, Carrick K. Dynamics of cervical remodeling during pregnancy and parturition: mechanisms and current concepts. Semin Reprod Med. 2007; 25:69–79. [PubMed: 17205425]
39. Word, RA.; Li, X-H. Cervical function during pregnancy and parturition. In: Petraglia, F.; Strauss, J.; Gabbe, SG.; Weiss, G., editors. Preterm Birth: Mechanisms, Mediators, Prediction & Interventions. Abingdon OX, UK: Taylor & Francis Medical Books; 2006.
40. Darios ES, Seitz B, Watts SW. Smooth muscle pharmacology in the isolated virgin and pregnant rat uterus and cervix. J Pharmacol Exp Ther. 2012; 341:587–596. [PubMed: 22366660]
41. Romero R, Espinoza J, Kusanovic JP, et al. The preterm parturition syndrome. BJOG. 2006; 113(Suppl 3):17–42. [PubMed: 17206962]
42. Romero R, Mazor M, Munoz H, Gomez R, Galasso M, Sherer DM. The preterm labor syndrome. Ann N Y Acad Sci. 1994; 734:414–429. [PubMed: 7978942]
43. Goldenberg RL, Gravett MG, Iams J, et al. The preterm birth syndrome: issues to consider in creating a classification system. Am J Obstet Gynecol. 2012; 206:113–118. [PubMed: 22177186]
44. Owen J, Iams JD, Hauth JC. Vaginal sonography and cervical incompetence. Am J Obstet Gynecol. 2003; 188:586–596. [PubMed: 12592276]
45. Iams JD, Johnson FF, Sonek J, Sachs L, Gebauer C, Samuels P. Cervical competence as a continuum: a study of ultrasonographic cervical length and obstetric performance. Am J Obstet Gynecol. 1995; 172:1097–1103. discussion 104–6. [PubMed: 7726247]
46. Harger JH. Cerclage and cervical insufficiency: an evidence-based analysis. Obstet Gynecol. 2002; 100:1313–1327. [PubMed: 12468179]
47. Caritis SN, Simhan H. Cervical pessary use and preterm birth: how little we know. Lancet. 2012; 379:1769–1770. [PubMed: 22475492]
48. Reiter E, Nielsen KA, Fedder J. Digital examination and transvaginal scan - competing or complementary for predicting preterm birth? Acta Obstet Gynecol Scand. 2012; 91:428–438. [PubMed: 22168831]

49. Berghella V, Baxter JK, Hendrix NW. Cervical assessment by ultrasound for preventing preterm delivery. Cochrane Database Syst Rev. 2013 CD007235.
50. Lim AC, Hegeman MA, Huis In TVMA, Opmeer BC, Bruinse HW, Mol BW. Cervical length measurement for the prediction of preterm birth in multiple pregnancies: a systematic review and bivariate meta-analysis. Ultrasound Obstet Gynecol. 2011; 38:10–17. [PubMed: 21465606]
51. Berghella V. Universal cervical length screening for prediction and prevention of preterm birth. Obstet Gynecol Surv. 2012; 67:653–658. [PubMed: 23112072]
52. Leitich H, Brunbauer M, Kaider A, Egarter C, Husslein P. Cervical length and dilatation of the internal cervical os detected by vaginal ultrasonography as markers for preterm delivery: A systematic review. Am J Obstet Gynecol. 1999; 181:1465–1472. [PubMed: 10601930]
53. Barros-Silva J, Pedrosa AC, Matias A. Sonographic measurement of cervical length as a predictor of preterm delivery: a systematic review. J Perinat Med. 2013:1–13.
54. Adriaensen WJ, Mathei C, Buntinx FJ, Arbyn M. A framework provided an outline toward the proper evaluation of potential screening strategies. J Clin Epidemiol. 2013; 66:639–647. [PubMed: 23395357]
55. 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; 174:990–995. [PubMed: 8633681]
56. Heath VC, Southall TR, Souka AP, Elisseou A, Nicolaides KH. Cervical length at 23 weeks of gestation: prediction of spontaneous preterm delivery. Ultrasound Obstet Gynecol. 1998; 12:312–317. [PubMed: 9819868]
57. To MS, Skentou CA, Royston P, Yu CK, Nicolaides KH. Prediction of patient-specific risk of early preterm delivery using maternal history and sonographic measurement of cervical length: a population-based prospective study. Ultrasound Obstet Gynecol. 2006; 27:362–367. [PubMed: 16565989]
58. Hassan SS, Romero R, Berry SM, et al. Patients with an ultrasonographic cervical length< or =15 mm have nearly a 50% risk of early spontaneous preterm delivery. Am J Obstet Gynecol. 2000; 182:1458–1467. [PubMed: 10871466]
59. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009; 62:1006–1012. [PubMed: 19631508]
60. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009; 151:264–269. W64. [PubMed: 19622511]
61. Bouyer J, Papiernik E, Dreyfus J, Collin D, Winisdoerffer B, Gueguen S. Maturation signs of the cervix and prediction of preterm birth. Obstet Gynecol. 1986; 68:209–214. [PubMed: 3737036]
62. Leung TN, Pang MW, Leung TY, Poon CF, Wong SM, Lau TK. Cervical length at 18–22 weeks of gestation for prediction of spontaneous preterm delivery in Hong Kong Chinese women. Ultrasound Obstet Gynecol. 2005; 26:713–717. [PubMed: 16308894]
63. Taipale P, Hiilesmaa V. Sonographic measurement of uterine cervix at 18–22 weeks' gestation and the risk of preterm delivery. Obstet Gynecol. 1998; 92:902–907. [PubMed: 9840546]
64. To MS, Skentou C, Liao AW, Cacho A, Nicolaides KH. Cervical length and funneling at 23 weeks of gestation in the prediction of spontaneous early preterm delivery. Ultrasound Obstet Gynecol. 2001; 18:200–203. [PubMed: 11555446]
65. Mortensen OA, Franklin J, Lofstrand T, Svanberg B. Prediction of preterm birth. Acta Obstet Gynecol Scand. 1987; 66:507–512. [PubMed: 3425252]
66. Fletcher, R.; Fletcher, SW. Diagnosis. In: Fletcher, R.; Fletcher, SW., editors. Clinical epidemiology: The essentials. Lippincott Williams & Wilkins; 2013.
67. Pires CR, Moron AF, Mattar R, Diniz AL, Andrade SG, Bussamra LC. Cervical gland area as an ultrasonographic marker for preterm delivery. Int J Gynaecol Obstet. 2006; 93:214–219. [PubMed: 16443225]
68. Owen J, Neely C, Northen A. Transperineal versus endovaginal ultrasonographic examination of the cervix in the midtrimester: a blinded comparison. Am J Obstet Gynecol. 1999; 181:780–783. [PubMed: 10521728]

69. Roh HJ, Ji YI, Jung CH, Jeon GH, Chun S, Cho HJ. Comparison of cervical lengths using transabdominal and transvaginal sonography in midpregnancy. J Ultrasound Med. 2013; 32:1721–1728. [PubMed: 24065252]
70. Mara M, Calda P, Haakova L, Zizka Z, Dohnalova A, Zivny J. Significance of ultrasound vaginal cervicometry in predicting preterm delivery. Med Sci Monit. 2002; 8:MT72–MT77. [PubMed: 12011784]
71. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003; 3:25. [PubMed: 14606960]
72. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. BMJ. 2003; 326:41–44. [PubMed: 12511463]
73. Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ. 2001; 323:157–162. [PubMed: 11463691]
74. Higgins, JP.; Green, S. Cochrane Handbook for Systematic Reviews of Interventions. West Sussex, England: The Cochrane Collaboration and John Wiley & Sons Ltd;
75. Hartmann K, Thorp JM, Mcdonald TL, Savitz DA, Granados JL. Cervical dimensions and risk of preterm birth: A prospective cohort study. Obstetrics and Gynecology. 1999; 93:504–509. [PubMed: 10214823]
76. De Carvalho MH, Bittar RE, Brizot Mde L, Bicudo C, Zugaib M. Prediction of preterm delivery in the second trimester. Obstet Gynecol. 2005; 105:532–536. [PubMed: 15738020]
77. Newman RB, Goldenberg RL, Iams JD, et al. Preterm prediction study: comparison of the cervical score and Bishop score for prediction of spontaneous preterm delivery. Obstet Gynecol. 2008; 112:508–515. [PubMed: 18757646]
78. To MS, Skentou C, Cicero S, Liao AW, Nicolaides KH. Cervical length at 23 weeks in triplets: prediction of spontaneous preterm delivery. Ultrasound Obstet Gynecol. 2000; 16:515–518. [PubMed: 11169343]
79. Mercer BM, Goldenberg RL, Das A, et al. The preterm prediction study: a clinical risk assessment system. Am J Obstet Gynecol. 1996; 174:1885–1893. discussion 93–5. [PubMed: 8678155]
80. Parra-Saavedra M, Gomez L, Barrero A, Parra G, Vergara F, Navarro E. Prediction of preterm birth using the cervical consistency index. Ultrasound Obstet Gynecol. 2011; 38:44–51. [PubMed: 21465603]
81. Hasegawa I, Tanaka K, Takahashi K, et al. Transvaginal ultrasonographic cervical assessment for the prediction of preterm delivery. J Matern Fetal Med. 1996; 5:305–309. [PubMed: 8972404]
82. Houlton MC, Marivate M, Philpott RH. Factors associated with preterm labour and changes in the cervix before labour in twin pregnancy. Br J Obstet Gynaecol. 1982; 89:190–194. [PubMed: 7066256]
83. Marivate M, De Villiers KQ, Fairbrother P. Effect of prophylactic outpatient administration of fenoterol on the time of onset of spontaneous labor and fetal growth rate in twin pregnancy. Am J Obstet Gynecol. 1977; 128:707–708. [PubMed: 327814]
84. Neilson JP, Verkuyl DA, Crowther CA, Bannerman C. Preterm labor in twin pregnancies: prediction by cervical assessment. Obstet Gynecol. 1988; 72:719–723. [PubMed: 3173923]
85. 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; 379:2162–2172. [PubMed: 22682464]
86. Heath VC, Southall TR, Souka AP, Novakov A, Nicolaides KH. Cervical length at 23 weeks of gestation: relation to demographic characteristics and previous obstetric history. Ultrasound Obstet Gynecol. 1998; 12:304–311. [PubMed: 9819867]
87. Brieger GM, Xie HN, Dawkins RR, et al. Transvaginal sonographic assessment of cervical dynamics during the third trimester of normal pregnancy. Acta Obstet Gynecol Scand. 1997; 76:118–122. [PubMed: 9049282]
88. Erasmus I, Nicolaou E, Van Gelderen CJ, Nicolaides KH. Cervical length at 23 weeks' gestation--relation to demographic characteristics and previous obstetric history in South African women. S Afr Med J. 2005; 95:691–695. [PubMed: 16327930]

89. Feltovich H, Hall TJ. Quantitative imaging of the cervix: setting the bar. Ultrasound Obstet Gynecol. 2013; 41:121–128. [PubMed: 23371343]

90. Berghella V, Owen J, Macpherson C, et al. Natural history of cervical funneling in women at high risk for spontaneous preterm birth. Obstet Gynecol. 2007; 109:863–869. [PubMed: 17400847]

91. Li Q, Zhang J, You Z. [Cervical ripening score by transperineal ultrasonography and its predictive effect for induction of labor by prostaglandin E2]. Chinese J Obstet Gynecol (in Chinese). 1998; 33:216–218.

92. Reeves MJ, Curtis CR, Salman MD, Stashak TS, Reif JS. Multivariable prediction model for the need for surgery in horses with colic. Am J Vet Res. 1991; 52:1903–1907. [PubMed: 1785737]

93. Pencina MJ, D'agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. Am J Epidemiol. 2012; 176:473–481. [PubMed: 22875755]

94. Morrow-Howell N. The M Word: Multicollinearity in multiple regression. Social Work Research. 1994; 18:247–262.

95. Salomon LJ, Diaz-Garcia C, Bernard JP, Ville Y. Reference range for cervical length throughout pregnancy: non-parametric LMS-based model applied to a large sample. Ultrasound Obstet Gynecol. 2009; 33:459–464. [PubMed: 19277949]

96. Berghella V, Roman A, Daskalakis C, Ness A, Baxter JK. Gestational age at cervical length measurement and incidence of preterm birth. Obstet Gynecol. 2007; 110:311–317. [PubMed: 17666605]

97. Owen J, Yost N, Berghella V, et al. Mid-trimester endovaginal sonography in women at high risk for spontaneous preterm birth. JAMA. 2001; 286:1340–1348. [PubMed: 11560539]

98. Guzman ER, Walters C, Ananth CV, et al. A comparison of sonographic cervical parameters in predicting spontaneous preterm birth in high-risk singleton gestations. Ultrasound Obstet Gynecol. 2001; 18:204–210. [PubMed: 11555447]

99. Feltovich H, Hall TJ, Berghella V. Beyond cervical length: emerging technologies for assessing the pregnant cervix. Am J Obstet Gynecol. 2012; 207:345–354. [PubMed: 22717270]

100. Goldenberg RL, Iams JD, Das A, et al. The Preterm Prediction Study: sequential cervical length and fetal fibronectin testing for the prediction of spontaneous preterm birth. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Am J Obstet Gynecol. 2000; 182:636–643. [PubMed: 10739521]

101. Espinoza J, Goncalves LF, Romero R, et al. The prevalence and clinical significance of amniotic fluid 'sludge' in patients with preterm labor and intact membranes. Ultrasound Obstet Gynecol. 2005; 25:346–352. [PubMed: 15789375]

102. Himaya E, Rhalmi N, Girard M, et al. Midtrimester intra-amniotic sludge and the risk of spontaneous preterm birth. Am J Perinatol. 2011; 28:815–820. [PubMed: 22094916]

103. Arabin B, Aardenburg R, Van Eyck J. Maternal position and ultrasonic cervical assessment in multiple pregnancy. Preliminary observations. J Reprod Med. 1997; 42:719–724. [PubMed: 9408871]

104. Arabin B, Roos C, Kollen B, Van Eyck J. Comparison of transvaginal sonography in recumbent and standing maternal positions to predict spontaneous preterm birth in singleton and twin pregnancies. Ultrasound Obstet Gynecol. 2006; 27:377–386. [PubMed: 16514620]

105. National Institutes of Health. [Retrieved July 20, 2012] Preterm Birth in Nulliparous Women: An Understudied Population at Great Risk (U10), 2008. Accessed Available from: http://grants.nih.gov/grants/guide/rfa-files/RFA-HD-08-029.html

106. Khoury MJ, Gwinn M, Ioannidis JP. The emergence of translational epidemiology: from scientific discovery to population health impact. Am J Epidemiol. 2010; 172:517–524. [PubMed: 20688899]

107. Reece EA. A clarion call for translational and collaborative research. Am J Obstet Gynecol. 2006; 194:1507–1509. [PubMed: 16731066]

108. Hiatt RA. Epidemiology: key to translational, team, and transdisciplinary science. Ann Epidemiol. 2008; 18:859–861. [PubMed: 18823793]

109. Viera AJ. Predisease: when does it make sense? Epidemiol Rev. 2011; 33:122–134. [PubMed: 21624963]

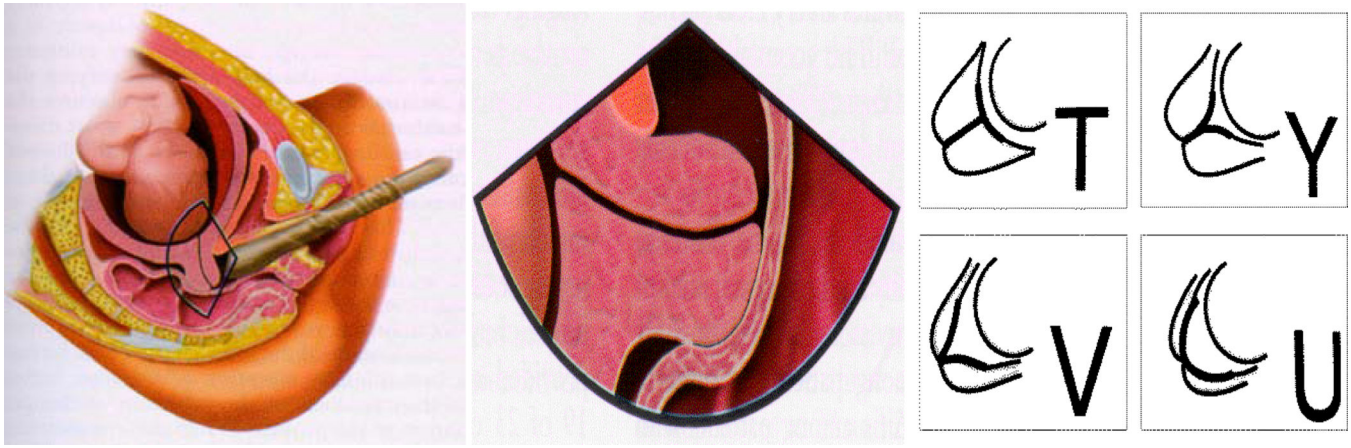

Figure 1.

Transvaginal ultrasonography (left) to assess a cervix (middle), [illustration by James A. Cooper, M.D., San Diego, CA in Callen (2008)[36]], showing cervical effacement and funneling (right)[37] [reproduced with permissions from Elsevier and the American Institute of Ultrasound in Medicine].

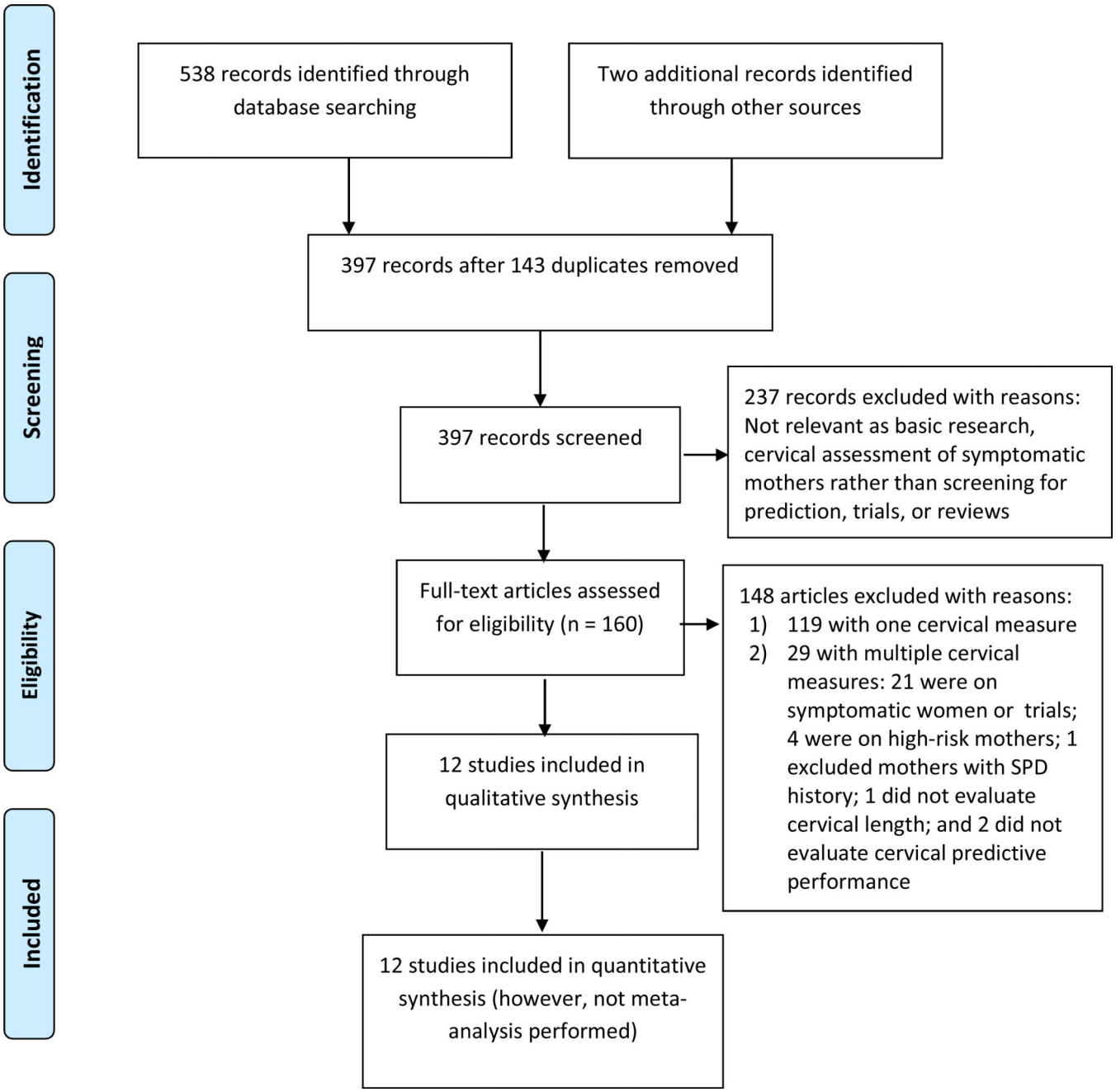

Figure 2.

Flow diagram of information through the different phases of the systematic review on comprehensive cervical screening for precocious cervical ripening to predict singleton preterm delivery in large observational cohort studies, 1980–2014.

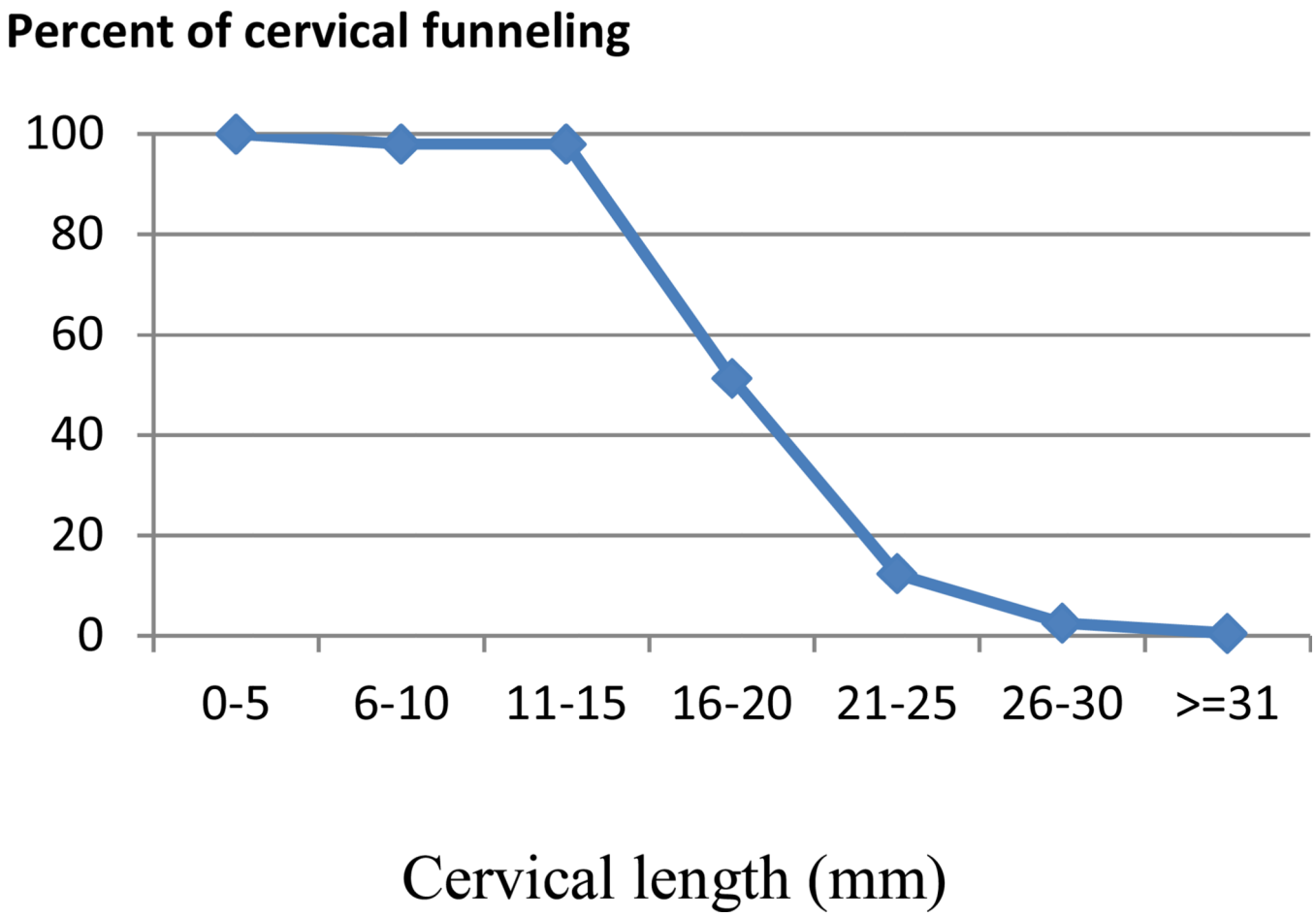

Figure 3.
Association of cervical length and percent of funneling based on To et al.[64]

Table 1

Characteristics of comprehensive cervical screening for precocious cervical ripening to predict singleton preterm delivery in large observational cohort studies, 1980–2014

Author, Published Year	Screen period	Location, venue	Population	Gestational weeks	Exam
Papiernik et al., 1986[27]	1971–1976	*Haguenau, France	4430	18, 19–24, 25–28, 29–31, 32–34, and 35–36	DE
Bouyer et al., 1986[61]	Ibid	Ibid	4390		DE
Mortensen et al,1987[65]	1982	*Skaraborg, Sweden	581	24, 28 & 32	DE
Hartmann et al,1999[75]	1995–2000	4 sites in North Carolina, USA	871	24–29	DE
Newman et al., 2008[77] Iams et al., 1996[8]	1992–1994 Ibid	10-site MFMU network, USA Ibid	2916, 2538 2915, 2531	22–24, 26–29 Ibid	DE & TVU DE & TVU
Hasegawa et al,1996[81]	1994	10 centers, Japan	729	15–34	TVU
Taipale & Hiilesmaa, 1998[63]	1995–1996	*Helsinki, Finland	3694	18–22	TVU
To et al., 2001[64]	1997–2000	*London, the UK	6334	22–24	TVU
de Carvalho et al, 2005[76]	1998–2001	São Paulo, Brazil	1958	21–24	TVU
Leung et al, 2005[62]	2000–2002	*Hong Kong, China	2880	18–22	TVU
Parra-Saavedra et al., 2011[80]	2009–2010	Barranquilla, Colombia	1115	5–36	TVU

Note: DE: digital examination; TVU: transvaginal sonography.

*for routine care. Ibid, the same as above; MFMU, NIH Maternal Fetal Medicine Unit Network in the U.S.

Table 2

Assessment of study quality

Author, Published Year	Consecutive sample	Prospective cohort design	Exam	Evaluators	Blinded	Reliability quantified	Outcome Defined	Predictive performance Assessed	Statistical association reported
Papiernik et al., 1986[27]	Yes	Yes	DE	OBs	Unclear	No	Yes	No	Yes
Bouyer et al., 1986[61]	Yes	Yes	DE	Ibid	Unclear	No	Yes	Yes	Yes
Mortensen et al,1987[65]	Yes	Yes	DE	Midwives	Unclear	No	Yes	Yes	Yes
Hartmann et al, 1999[75]	Yes	Yes	DE	OBs & nurses	Provider	No	Yes	Yes	No
Newman et al., 2008[77]	Yes	Yes	DE	Nurses & examiners	Provider	No	Yes	Yes	No
Iams et al., 1996[8]	Yes	Yes	DE & TVU	Nurses & examiners	Provider	No	Yes	Yes	Yes
Hasegawa et al, 1996[81]	Yes	Yes	TVU	?	Unclear	No	Yes	Yes	Yes
Taipale & Hiilesmaa, 1998[63]	Yes	Yes	TVU	OBs and midwives	Provider	No	Yes	Yes	Yes
To et al., 2001[64]	Yes	Yes	TVU	OBs & sonographers	Provider	No	Yes	No	Yes
de Carvalho et al, 2005[76]	Yes	Yes	TVU	Sonographers	Patient & Provider	No	Yes	Yes	No
Leung et al, 2005[62]	Yes	Yes	TVU	Sonographers	Patient & Provider	No	Yes	Yes	No
Parra-Saavedra et al., 2011[80]	Yes	Yes	TVU	OBs	Unclear	Yes	Yes	Yes	No

Note: OB, obstetrician; DE: digital examination; TVU: transvaginal sonography; ?: not clear.

Table 3

Measurements and performance of screening for precocious cervical ripening to predict singleton preterm delivery in large cohort studies, 1980–2014

Author, Published Year	Gestational weeks at initial assessment	Measurements (cm for CL)	Funneling Dilation %	Preterm (<wks)	Incidence of PTD %	Sensitivity %	Specificity %	PPV %	NPV %	ROC	LR+	LR−	Association
Papiernik et al., 1986[27]	18	Dilation, Station, Length, Uterine contraction, Expanded lower segment	0.8–12.4 By ≤18, 24, 28, 31, 34, 36 wks	37	?								2.5–3.4 (vary by 0.8–4.3 weeks)
													0.9–2.9
													1.2–2.9
													0.6–1.9
Bouyer et al., 1986[61]	18	Score: short open cervix, contraction, parity, age	1.1–14	37	5.9*	44–57*	71–78*				1.5–2.4*	0.59–0.78*	0.8–3.1/1.8–6.7*
					5.5	56–64	73–78				2.1–2.8	0.47–0.58	0.9–2.7/1.6–3.5
Mortensen et al,1987[65]	24	Modified Bishop score	4.0	37	1.5	33	88	4	97		2.8	0.76	
		Dilation				11	97	6	99		3.6	0.92	8.3
		Effacement											3.1
Hartmann et al,1999[75]	24–29	CL<2.0	6.0	37 or pPROM	8.3	**13**	93	15	92		1.9	0.94	
		Dilation ≥1.0				8	99	38	92		8	0.93	
		Cervical Score^<2				**20**	93	21	92		**2.9**	0.86	
Newman et al., 2008[77]	22–24	T1: Bishop score ≥4	?	35	4.4	28	90.9	12		0.66	3.0	0.80	
		Cervical score^<1.5				13	97.9	21		0.61	6.4	0.88	
		T2: Bishop score ≥5	?			32	93.0	14		0.68	4.6	0.73	
		TVU CL<2.0				**32**	95	17	98	0.68	6.4	0.72	
		TVU CF present				32	91	11	98		3.6	0.75	
		Cervical score^<1.5				**36**	95	20	98		**7.4**	0.67	
Iams et al., 1996[8]	15–34	T1: Bishop≥4	6.4	35	4.3	28	91	12	97		3.0	0.80	
		CL ≤2.5				37	92	18	97		4.8	0.68	6.19
		CF present				25	95	17	97		4.6	0.79	
		T2: Bishop≥4	9.1			43	83	10	97		2.4	0.70	
		CL ≤2.5				49	87	11	98		3.7	0.58	9.57

Author, Published Year	Gestational weeks at initial assessment	Measurements (cm for CL)	Funneling Dilation %	Preterm (<wks)	Incidence of PTD %	Sensitivity %	Specificity %	PPV %	NPV %	ROC	LR+	LR−	Association
		CF present				33	92	17	98		3.9	0.34	
Hasegawa et al,1996[61]	15–34	CL ≤2.7	7.8	36	3.3			10*/2					4.86 (1.85–12.72)*
		Open internal os. Funneling index						7*/11					6.00 (1.65–21.71)
Taipale & Hiilesmaa, 1998[63]	18–22	CL ≤2.9	0.7	35	0.8	**19**	97	6			**6.3**	0.84	8
		Dilation ≥0.5		37	2.4	16	99	20			16	0.85	28
		Either				**29**	97	7			**9.7**	0.73	11
To et al., 2001[64]	22–24	CL	4	33	0.9								24.9
		Internal os. ≥0.5											1.8
de Carvalho et al, 2005[76]	21–24	CL ≤2	1.5	34	3.4	**7**							
		Add CF present				**34**							
Leung et al, 2005[62]	18–22	**CL ≤2.7**	6.3 width 6.4 depth 4.3	34	0.7	**37**	96	6	100		**9.8**	0.66	
		CF				32	94	3	100		5.2	0.73	
		Both				26	99	15	100		**26**	0.74	
		Either				**42**	91	3	100		4.7	0.64	
Parra-Saavedra et al., 2011[80]	5–36	**Consistency index**	Excluded	34	2.1	**64**	98	47	99	0.94	**39.7**		
		CL				9	98	9	98		4.3		

Note: The bolded composite measures had better predictive performance than cervical length alone. T1 and T2, assessments at two time points.

*for nulliparous women, other values in the same cell for parous women;

^Cervical Score= Cervical length (cm) – Cervical dilation (cm). CL: cervical length; CF: cervical funneling; DE: digital examination; TVU: transvaginal ultrasonography; PTD, preterm delivery; wks: weeks; Ibid: the same as above; ROC: receiver operating characteristic curves; PPV: positive predictive value; NPV: negative predictive value; LR: likelihood ratio, calculated results based on original values in papers before being rounded.

Table 4

Indicators and performance of screening for composite measures of precocious cervical ripening in selected large observational cohort studies, 1980–2014

Mode of assessment	Study, Year	Time (wks)	Indicators # (cm for CL)	PTD (wks)	Prevalence	Sensitivity %	Specificity %	+PV %	−PV %	+LR
I: DE only	Hartmann et al,1999[75]	24–29	**CL<2.0**	37	8.3	**13**	93	15	92	**1.9**
			Dilatation ≥1.0			8	99	38	92	8
			Score*>2			**20**	93	21	92	**2.9**
			Difference			**+7**				**+1**
II: DE and TVU	Newman et al., 2008[77]	26–29	**CL ≤2.0**	35	4.3	**32**	95	17	98	**6.4**
		T2	Funneling			32	91	11	98	3.5
			Score<1.5			**36**	95	20	98	**7.4**
			Difference			**+4**				**+1**
	Taipale and Hiilesmaa, 1998[63]	18–22	**CL ≤2.9**	35	0.8	**19**	97	6		**6.3**
			Dilatation≥0.5	37	2.4	16	99	20		16
			Either			**29**	97	7		**9.7**
			Difference			**+10**				**+3.4**
III: TVU	de Carvalho et al,2005[76]	21–24	**CL ≤2**	34	3.4	**7**				
			CL+CF			**34**				
			Difference			**+27**				
	Leung et al, 2005[62]	18–22	**CL ≤2.7**	34	0.7	**37**	96	6	100	**9.8**
			Funneling			32	94	3	100	5.2
			Both			26	99	15	100	**26**
			Either			**42**	91	3	100	4.7
			Difference			**+5**				**+16.2**

Note. CL, cervical length; DE: digital examination; Difference: performance difference between the composite cervical assessment and cervicallength only; PTD, preterm delivery; TVU transvaginal ultrasonography;

*Score=Cervical length-Dilatation; PV, predictive value; LR, likelihood ratio; wks: weeks.