OBSTETRICS

Predictive accuracy of changes in transvaginal sonographic cervical length over time for preterm birth: a systematic review and metaanalysis

Q6 Agustin Conde-Agudelo, MD, MPH, PhD; Roberto Romero, MD, DMedSci

The objective of the study was to determine the accuracy of changes in transvaginal sonographic cervical length over time in predicting preterm birth in women with singleton and twin gestations. Sources of data included the following: PubMed, Embase, Cinahl, Lilacs, and Medion (all from inception to May 31, 2015), bibliographies, Google scholar, and conference proceedings. This study included cohort or cross-sectional studies reporting on the predictive accuracy for preterm birth of changes in cervical length over time. Two reviewers independently selected studies, assessed the risk of bias, and extracted the data. Summary receiver-operating characteristic curves, pooled sensitivities, specificities, and likelihood ratios were generated. Fourteen studies met the inclusion criteria, of which 7 provided data on singleton gestations (3374 women) and 8 on twin gestations (1024 women). Among women with singleton gestations, the shortening of cervical length over time had a low predictive accuracy for preterm birth at less than 37 and less than 35 weeks of gestation with pooled sensitivities, specificities, and positive and negative likelihood ratios ranging from 49% to 74%, 44% to 85%, 1.3 to 4.1, and 0.3 to 0.7, respectively. In women with twin gestations, the shortening of cervical length over time had a low to moderate predictive accuracy for preterm birth at less than 34, less than 32, less than 30, and less than 28 weeks of gestation with pooled sensitivities, specificities, and positive and negative likelihood ratios ranging from 47% to 73%, 84% to 89%, 3.8 to 5.3, and 0.3 to 0.6, respectively. There were no statistically significant differences between the predictive accuracies for preterm birth of cervical length shortening over time and the single initial and/or final cervical length measurement in 8 of 11 studies that provided data for making these comparisons. In the largest and highest-quality study, a single measurement of cervical length obtained at 24 or 28 weeks of gestation was significantly more predictive of preterm birth than any decrease in cervical length between these gestational ages. Change in transvaginal sonographic cervical length over time is not a clinically useful test to predict preterm birth in women with singleton or twin gestations. A single cervical length measurement obtained between 18 and 24 weeks of gestation appears to be a better test to predict preterm birth than changes in cervical length over time.

Key words: longitudinal studies, prematurity, predictive value of test, screening, shortening in cervical length, singleton gestation, twin gestation

Preterm birth is one of the great obstetrical syndromes, which are characterized by multiple etiologies, a long preclinical stage, frequent fetal involvement, clinical manifestations that are often adaptive in nature, and complex interactions between the fetal and maternal genome and the environment that may predispose to the syndrome.[1,2]

Transvaginal sonographic measurement of cervical length (CL) provides useful information about one of the mechanisms of disease implicated in the etiology of the preterm parturition syndrome. In 1990, Andersen et al[3] published a seminal study in which a transvaginal sonographic CL below the 50th percentile at 30 weeks of gestation was associated with a 3.7-fold increased risk of preterm birth compared with a CL at or above the 50th percentile. Logistic regression analysis showed a progressive and statistically significant trend toward a higher risk of preterm birth with a shorter CL. Moreover, it was reported that a CL less than 39 mm had a sensitivity of 76% and a specificity of 59% to predict preterm birth at less than 37 weeks of gestation. Since then, transvaginal sonographic CL has been extensively investigated as a predictor of preterm birth.[4-11]

Several metaanalyses have now provided compelling evidence that a transvaginal sonographic CL measurement at 18–24 weeks of gestation is one of

Q1 From the Perinatology Research Branch, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development/National Institutes of Health/Department of Health and Human Services, Bethesda, MD, and Detroit, MI (Drs Conde-Agudelo and Romero), Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, MI (Dr Romero), Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI (Dr Romero), and Department of Molecular Obstetrics and Genetics, Wayne State University, Detroit, MI (Dr Romero).

Received April 28, 2015; revised May 29, 2015; accepted June 4, 2015.

This study was supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services.

The authors report no conflict of interest.

Corresponding author: Roberto Romero, MD, DmedSci. romeror@mail.nih.gov

0002-9378/$36.00
Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.ajog.2015.06.015

REV 5.2.0 DTD ■ YMOB10462_proof ■ 7 July 2015 ■ 3:36 pm ■ ce

the strongest and most consistent predictors of preterm birth in asymptomatic women with singleton gestations, regardless of whether they have a history of preterm birth,[12-15] and twin gestations.[16-18]

More recently, an analysis of serial measurements of transvaginal sonographic CL has shown that assessment of risk for preterm birth can be further refined. Several studies have reported that the shortening of transvaginal sonographic CL over time is associated with an increased risk of preterm birth,[9,19-23] whereas other studies have not been able to demonstrate this association.[24-26] Recently there has been a renewed interest in the relationship between CL changes over time and the risk of preterm birth.[27] The shortening of transvaginal sonographic CL over time has been proposed as a better predictor of spontaneous preterm birth than a single CL measurement.[9] However, to the best of our knowledge, there are no studies that have systematically evaluated the predictive performance of this test.

The primary aim of this study was to determine the accuracy of changes in transvaginal sonographic CL over time to predict preterm birth in women with singleton and twin gestations through the use of formal methods for systematic reviews and metaanalytic techniques.

Materials and methods

This study followed a prospectively prepared protocol and is reported in accordance with recommended methods for systematic reviews of diagnostic test accuracy.[28,29] The 2 authors independently retrieved and reviewed studies for eligibility, assessed their risk of bias, and extracted data. All disagreements encountered in the review process were resolved through consensus.

Data sources and searches

To identify potentially eligible studies, we searched PubMed, Embase, Cinahl, Lilacs, and Medion (all from inception to May 31, 2015) using an existing literature search strategy for systematic reviews of predictive tests for preterm birth.[30] Google Scholar, proceedings of congresses on preterm birth, ultrasound in obstetrics and maternal-perinatal medicine, and reference lists of identified studies were also searched. No language restrictions were applied.

Eligibility criteria

The systematic review focused on cohort or cross-sectional studies that reported on the accuracy of changes in transvaginal sonographic CL over time to predict preterm birth in asymptomatic pregnant women with a singleton or twin gestation, and that allowed a construction of 2×2 contingency tables.

Studies were excluded if they had the following characteristics: (1) were case-control studies because there is consistent evidence that they are associated with higher diagnostic or predictive accuracy compared with cohort studies[31]; (2) assessed CL changes over time in women with cervical cerclage or pessary, preterm labor, premature rupture of membranes, or those who were receiving progestogens; (3) were reviews, case series or reports, editorials, or letters without original data; or (4) did not publish accuracy test estimates and sufficient information to calculate them could not be retrieved. For studies that resulted in multiple publications, the data from the one with the largest sample size were used and supplemented if additional information appeared in the others.

Reference standard outcomes

The reference standard outcomes included the following: in women with singleton gestations, spontaneous preterm birth at less than 37 and less than 35 weeks of gestation; in women with twin gestations, spontaneous preterm birth at less than 34, less than 32, less than 30, and less than 28 weeks of gestation.

Assessment of risk of bias

Study quality was assessed using a modified version of the Quality Assessment of Diagnostic Accuracy Studies-2 tool.[32] The assessments were judged as low risk, high risk, or unclear risk of bias. The items were evaluated and interpreted as follows:

1. Patient selection. Low risk of bias: women were consecutively or randomly selected; high risk of bias: convenience sampling (arbitrary recruitment or nonconsecutive recruitment).
2. Description of the test. Low risk of bias: the study described sufficient details of the technique used for measuring CL such as the plane in which images were obtained, anatomic references for the determination of CL, and number of measurements; high risk of bias: if this information was not reported.
3. Reference standard. Low risk of bias: spontaneous preterm birth, defined as a preterm delivery after the spontaneous onset of contractions or preterm premature rupture of membranes, regardless of whether the delivery was vaginal, by cesarean delivery, or, in the case of rupture of membranes, induced; high risk of bias: inclusion of both spontaneous and indicated preterm birth in the reference standard.
4. Blinding. Low risk of bias: the study clearly stated that clinicians managing the patient did not have knowledge of the CL measurement results; high risk of bias: unmasking of clinicians to test results.
5. Inclusion of women in the analysis. Low risk of bias: if at least 90% of women recruited into the study were included in the analysis; high risk of bias: if less than 90% of women recruited into the study were included in the analysis.
6. Use of interventions aimed to prevent preterm birth based on the test results. Low risk of bias: clinicians did not use interventions based on the results of the CL measurements; high risk of bias: clinicians used interventions based on the results of the test (eg, cerclage, pessary, vaginal progesterone).

If there was insufficient information available to make a judgment about these items, then they were scored as unclear risk of bias. We did not calculate a summary score estimating the overall quality of each study because of the well-known problems associated with such scores.[33]