Accuracy of cervical transvaginal sonography in predicting preterm birth: a systematic review


CRD summary
This review assessed the diagnostic accuracy of transvaginal cervical sonography in predicting spontaneous pre-term birth. The authors concluded that cervical length and/or funneling can help predict spontaneous pre-term birth in asymptomatic women. Although there were differences among the studies and the diagnostic accuracy depended upon definitions of abnormality, the authors' general conclusions are likely to be reliable.

Authors' objectives
To assess the diagnostic accuracy of transvaginal cervical sonography in predicting spontaneous pre-term birth.

Searching
MEDLINE (1966 to June 2002), EMBASE (1980 to June 2002), Pascal (1973 to June 2002), BIOSIS Previews (1969 to June 2002), the Cochrane Library (Issue 2, 2002), MEDION (1974 to December 2000), the National Research Register (Issue 4, 2004), SciSearch (1974 to June 2002) and Conference Papers Index (1973 to June 2002) were searched for reports published in any language. A reference giving details of the electronic search strategy was provided (see Other Publications of Related Interest). The reference lists of identified studies and reviews were also checked.

Study selection
Studies of observational cohort design were eligible for inclusion, whereas case-control studies were excluded.

Specific interventions included in the review
Studies of antenatal transvaginal sonographic cervical length measurement of the cervix were eligible for inclusion. The included studies measured cervical length at between 7 and 30 weeks' gestation, with most studies measuring the length at between 20 and 24 weeks. The most common cut-off point for abnormality was 25 cm and the test thresholds ranged from 10 to 60 mm. The included studies also assessed the presence of funneling.

Reference standard test against which the new test was compared
The reference standard was gestational age at spontaneous pre-term birth. The specific definitions of pre-term used in the included studies were 32, 34 and 37 weeks' gestation. The review focused on birth before 34 weeks and included studies with spontaneous birth before 33 weeks and before 35 weeks in this category.

Participants included in the review
Studies of asymptomatic or symptomatic pregnant women were eligible for inclusion if the gestation at spontaneous birth was known. The review defined asymptomatic as without uterine tightenings or contractions, and symptomatic as with uterine tightenings or contractions but before advanced cervical dilatation.

Outcomes assessed in the review
The studies included in the review assessed diagnostic accuracy. To be included in the analysis, the studies had to present data separately for singleton and twin pregnancies.

How were decisions on the relevance of primary studies made?
Two reviewers independently screened the results of the literature search. Two reviewers independently selected English language publications, one reviewer selected French language reports, and an unidentified number of people who had sufficient command of the language to extract data selected reports in other languages. Any disagreements were resolved by consensus or through discussion with a third reviewer.
Assessment of study quality
Studies were considered to be of high quality if they used a prospective design, enrolled consecutive women, provided an adequate description of the test and included a blinded assessment of the test result. The authors did not state how many reviewers performed the quality assessment.

Data extraction
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

Data were extracted separately for singleton and twin pregnancies, for different thresholds defining abnormality, and for three groups of gestational age at testing (less than 20 weeks, 20 to 24 weeks and greater than 24 weeks’ gestation). The gestational age group was determined by the mean gestational age at testing. Data were extracted to construct 2x2 tables for various cervical lengths, funneling and gestational age at pre-term birth. For each study, the likelihood ratios (LRs) and 95% confidence intervals (CIs) were calculated according to gestational age at testing, thresholds defining abnormalities, and the three reference standards used. Authors of studies that did not present sufficient information to allow the calculation of accuracy data were contacted for relevant data.

Methods of synthesis
How were the studies combined?
The studies were combined where possible in a meta-analysis. The data were combined separately for studies of asymptomatic women with singleton and twin pregnancies, and for symptomatic women with singleton pregnancies. Subgroup analyses were performed according to gestational age at testing, thresholds defining abnormalities, and the three reference standards used (32, 34 and 37 weeks). Summary LRs and 95% CIs for subgroups were calculated where more than one study provided data. The areas under summary receiver operating characteristic (ROC) curves was calculated and used as a measure of the overall accuracy of the various studies. Publication bias was assessed using a funnel plot.

How were differences between studies investigated?
Heterogeneity was assessed separately for studies of singleton and twin pregnancies by plotting data for the subgroups using ROC curves. In addition, pooled LRs for the diagnostic accuracy of cervical sonography and funneling were calculated for high-quality studies.

Results of the review
Forty-six studies (31,577 women) were included. Of these, 33 studies were in asymptomatic women and 13 studies were in symptomatic women.

Eight studies met all four criteria for high quality: 6 studies of asymptomatic women with singleton pregnancies and 2 studies of asymptomatic women with singleton pregnancies.

Asymptomatic women.
The LRs for positive and negative tests varied according to the cut-offs and gestational age at testing. The area under the summary ROC curve was 0.80 (95% CI: 0.69, 0.91) for singleton pregnancies and 0.67 (95% CI: 0.55, 0.79) for twin pregnancies.

Spontaneous pre-term birth was more accurately predicted by studies testing before 20 weeks and using smaller cervical lengths for both singleton and twin pregnancies.

With the exception of one subgroup (one twin pregnancy subgroup), no heterogeneity was found in the LRs of positive tests for all the subgroups tested. Heterogeneity was detected in some LRs for negative tests.

For asymptomatic women of less than 20 weeks’ gestation, when using a cervical threshold of 25 mm and a reference standard of pre-term delivery before 34 weeks (5 studies), the LR for a positive test was 6.29 (95% CI: 3.29, 12.02).
This would render a 4.1% pre-test probability into a 15.8% post-test probability. The corresponding LR for a negative test was 0.79 (95% CI: 0.65, 0.95).

Studies showed that the larger the funneling, the more accurately it predicted pre-term birth.

High-quality studies showed similar results to all studies combined.

Symptomatic women.

There were few studies assessing the diagnostic accuracy of cervical sonography.

The LRs for positive and negative tests varied according to the cut-offs and gestational age at testing.

Studies showed that funneling appeared to predict spontaneous pre-term birth.

The funnel plot showed no evidence of publication bias.

**Authors' conclusions**

Both cervical length and funneling, either alone or in combination, can help in predicting spontaneous pre-term birth in asymptomatic women. There were limited data on symptomatic women, but the findings showed that funneling appeared to predict spontaneous pre-term birth.

**CRD commentary**

The review question was clear in terms of the study design, test and reference standard, and participants. Several relevant sources were searched and no language restrictions were applied. Two reviewers independently selected English language publications and this reduced the potential for bias and errors. However, the methods used to assess quality and extract the data were not described; hence, any efforts made to reduce errors and bias cannot be judged. Quality was assessed using established criteria and the results were reported.

Adequate information on the included studies was tabulated and appropriate statistical methods were used to synthesise the data. Various subgroup analyses were undertaken and the results for high-quality studies were analysed separately. The authors discussed some of the limitations of the review, such as the need for caution when interpreting LRs in the presence of heterogeneity and the limited power due to the small number of studies within each subgroup; many subgroups only contained one study. The authors’ general conclusions are likely to be reliable. However, as the authors pointed out, the usefulness of cervical length measurement and funneling depends upon the method chosen to define the abnormality. Furthermore the clinical heterogeneity among the included studies must be borne in mind.

**Implications of the review for practice and research**

Practice: The authors stated that with the provision of individual threshold estimates stratified according to common testing gestational ages, clinicians should be able to use the more realistic probabilities presented in the review to predict spontaneous pre-term birth.

Research: The authors did not state any implications for further research.

**Funding**

WellBeing, charity grant number K2/00.

**Bibliographic details**

PubMedID
12942506

DOI
10.1002/uog.202

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Data Collection; Female; Humans; Obstetric Labor, Premature /ultrasonography; Pregnancy; Prospective Studies; Sensitivity and Specificity; Ultrasonography, Prenatal /methods /standards

AccessionNumber
12003001971

Date bibliographic record published
31/03/2005

Date abstract record published
31/03/2005

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.