Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://hdl.handle.net/11245/2.107568

File ID	uvapub:107568
Filename	Chapter 7: Cervical length measurement for the prediction of preterm birth in multiple pregnancies: a systematic review and bivariate meta-analysis
Version	unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	PhD thesis
Title	Progesterone for the prevention of preterm birth
Author(s)	A.C. Lim
Faculty	AMC-UvA
Year	2012

FULL BIBLIOGRAPHIC DETAILS:
http://hdl.handle.net/11245/1.385737

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content licence (like Creative Commons).

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
(pagedate: 2014-11-23)

7

Cervical length measurement for the prediction of preterm birth in multiple pregnancies: a systematic review and bivariate meta-analysis

Arianne C. Lim, Maud A. Hegeman,
Maite A. Huis in 't Veld, Brent C. Opmeer,
Hein W. Bruinse and Ben Willem J. Mol

Ultrasound in Obstetrics and Gynecology. 2011;
38: 10–17

Abstract

Objectives To review the literature on cervical length as a predictor of preterm birth in asymptomatic women with a multiple pregnancy.

Methods We searched MEDLINE, Embase and reference lists of included articles to identify all studies that reported on the accuracy of cervical length for predicting preterm birth in asymptomatic women with a multiple pregnancy. We scored study characteristics and study quality, and extracted data in order to construct two-by two tables cross-classifying cervical length and preterm delivery. Meta-analysis using a bivariate model was performed. Summary receiver–operating characteristics (ROC) curves were generated for various test characteristics and outcome definitions.

Results We found 21 studies reporting on 2757 women. There was a large variation in gestational age at measurement, cut-off point for cervical length and definition of preterm birth. The summary ROC curve indicated a good predictive capacity of short cervical length for preterm birth. Summary estimates of sensitivity and specificity for preterm birth before 34 weeks' gestation were 78% and 66%, respectively, for 35 mm, 41% and 87% for 30 mm, 36% and 94% for 25 mm and 30% and 94% for 20 mm.

Conclusions In women with a multiple pregnancy, second-trimester cervical length is a strong predictor of preterm birth. In the absence of effective preventive strategies, there is currently no place in clinical practice for cervical length measurement in this population. However, future studies should evaluate preventive interventions in women with multiple pregnancies and a short cervix, and cervical length should be measured in any trial studying preventive strategies in multiple pregnancies.

Introduction

Despite the enormous advances in neonatal care during recent years, preterm birth remains the major cause of handicap in children without congenital anomalies or genetic disorders. Many studies have been conducted into the clinical relevance of cervical length as a predictor of preterm birth. In both symptomatic and asymptomatic pregnant women, it has been shown that transvaginal sonography of the cervix is able to identify women who are at increased risk for preterm birth.[1] This finding has led to randomized controlled trials in which women with asymptomatic cervical shortening during the second trimester were treated with progesterone or placebo. In the first trial that was reported, daily vaginal progesterone turned out to reduce the risk of preterm birth by 40%.[2] Similar results with progesterone had previously been obtained in women with a singleton pregnancy and a history of preterm birth.[3, 4] In unselected women with multiple pregnancies, progesterone has appeared to be ineffective in reducing preterm birth rates so far.[5, 6]

The mechanism of preterm birth in women with a multiple pregnancy is likely to differ from that in women with singleton pregnancies. The human body often seems to lack the natural capacity to carry multiple fetuses to term, whereas the cause of preterm birth in singleton pregnancies is more likely to be found in individual maternal or fetal factors. However, it might be possible that these individual factors also play a role in a minority of women with a multiple pregnancy. Consequently, tests used to predict preterm birth in women with a singleton pregnancy may be useful in women with a multiple pregnancy, although the relative over-distension of the uterus in multiple pregnancies might affect the cervix and therefore the predictive value of cervical length measurement.

We conducted a systematic review to assess whether sonographically measured cervical length can be used to predict preterm birth in multiple pregnancies. We limited our search to asymptomatic women in the second trimester. We planned a meta-analysis using bivariate regression analysis, accounting for correlation between sensitivity and specificity.

Methods

Electronic searches

With the help of a librarian, we carried out electronic searches of MEDLINE and Embase from inception (1966 and 1947, respectively) to January 2009. The search strategy consisted of MeSH or key terms related to multiple pregnancy, cervix and preterm birth. We checked reference lists of relevant studies to identify cited articles not captured by electronic searches and contacted authors of primary studies where contact addresses were available. We aimed to identify studies that reported on cervical length as a predictor of preterm birth in asymptomatic women with a multiple pregnancy. We did not use any restricting criteria for the study design. Reference Manager 11.0 (Thomson ISI ResearchSoft, Carlsbad, CA, USA) databases were established incorporating results of all searches.

Selection of studies

Articles identified through the initial search were first screened on title and, where available, on abstract by two independent reviewers (A.L. and M.H.V.). Studies were included if they reported on women with a multiple pregnancy in whom cervical length was sonographically measured during pregnancy, and for whom gestational age at birth was known. When studies could not be excluded based on their title or abstract, a full manuscript was obtained. These were then independently scored by two reviewers (A.L. and M.H.). When a manuscript was written in a language other than English or Dutch, the article was translated by a colleague with both fluency in the language and expertise in the subject area. At this stage a fourth inclusion criterion was added: the reviewer had

to be able to subtract a two-by-two table cross-classifying cervical length and gestational age from the article. In cases where this was not possible, but where the results that were presented indicated that the original data would allow for the generation of a two-by-two table, the authors were contacted by e mail and/or post. We included studies examining pregnant women at any level of risk and in any healthcare setting. Any disagreements were resolved by consensus or a third reviewer (B.M.). No language restrictions were applied.

Data extraction

For all articles that were included, two reviewers (A.L. and M.H.) independently extracted descriptive data (first author and year of publication), study design characteristics, test characteristics (transvaginal, transperineal or abdominal ultrasound and cut-off values for cervical length and gestational age at measurement), definition of outcome (cut-off values for preterm birth) and data on study quality and test accuracy, using a pre-designed and piloted data extraction form.

The two reviewers independently assessed all included manuscripts for study quality according to an adapted version of QUADAS.[7] Studies were classified as having an adequate description of the testing procedure when at least the following two aspects were mentioned: (1) the bladder was emptied before testing; (2) whether or not funneling was included in the cervical length.

Data synthesis

We calculated sensitivity and specificity with 95% confidence intervals (CIs) for each study individually, created forest plots to explore heterogeneity for sensitivity and specificity and plotted their combined results in receiver–operating characteristics (ROC) space (sensitivity vs. 1 − specificity). Bivariate regression analysis was used to obtain summary estimates of sensitivity and specificity and their 95% CIs, and to construct summary ROC (sROC) curves.[8] With a bivariate regression model summary estimates can be calculated simultaneously for sensitivity and specificity within a single model. Sensitivity and specificity within a study are often negatively correlated, owing to implicit variation of threshold values. The bivariate regression model statistically incorporates the negative correlation that might exist between sensitivity and specificity. Variation or heterogeneity between the results of the studies included in the meta-analysis can be the result of differences in thresholds, but also of chance, bias due to flawed design, different clinical subgroups and unexplained variation. When necessary, the bivariate model uses a random effects approach, due to which clinical heterogeneity beyond chance is accounted for.

Different studies often reported a different threshold value to define a positive test result for cervical length, and several studies also reported accuracy for multiple threshold values. In addition, studies used different threshold values for gestational age at delivery to define preterm birth. If we estimated accuracy for a single combination of threshold values for cervical length and preterm birth, this estimate would be based on only a limited number of studies. Furthermore, it is not clear which is the appropriate threshold

for either definition. In order to evaluate accuracy measures over the whole range of possible thresholds, however, we did not limit our analysis to a single threshold value, but estimated accuracy measures for all reported threshold values by assuming that the shift in accuracy (higher sensitivity and lower specificity) due to different thresholds is accounted for by the correlation term, as specified in the bivariate model.

As the recommended bivariate modeling approach cannot appropriately account for covariates (such as threshold) with multiple observations from the same study, we used the following strategy to statistically incorporate differences in cervical length thresholds and definitions of preterm birth in the model. We performed three types of analysis, and integrated their results into summary estimates for different cervical length and preterm birth threshold combinations. The first type includes all reported accuracy estimates, irrespective of threshold values for cervical length and preterm birth or gestational age at which cervical length was measured; the second type analyzes the bivariate model for four different cervical length thresholds (20, 25, 30 and 35 mm); and the third type fitted the bivariate model for three preterm birth thresholds (29, 34 and 37 weeks).

In order to avoid results being biased towards studies reporting on many different thresholds, we estimated each model in 100 stratified bootstrap samples, in which only one accuracy estimate from each study was randomly selected. For each parameter the average overall estimates from 100 bootstrap samples is reported. The results of the model were used to estimate summary ROC curves, where the increase in sensitivity and decrease in specificity reflect the shift in threshold value of cervical length in the model, resulting in separate ROC curves for different criteria to define preterm birth (<34 weeks, >34 weeks). In a subgroup analysis, we also performed separate analyses for twins and triplets.

Results

Figure 1 summarizes the flow of studies through the review. Out of 382 potentially relevant abstracts, 21 studies screening 2757 women were included. Authors of a further 13 studies were contacted about unclear information in their articles. None of the contacted authors provided information that could clarify our uncertainties, the main reason given being that 'data were no longer available'. These studies were therefore not included in our analysis.

Twelve studies were prospective and nine studies were retrospective cohort studies. Sample sizes ranged from 14 to 383 women. Table 1 shows the study characteristics of all included studies. Most studies reported on several measurement times, cut offs for cervical length and cutoffs for preterm birth, allowing for multiple two-by-two tables to be extracted from one study. The number of two-by-two tables per study ranged from 1 to 20.

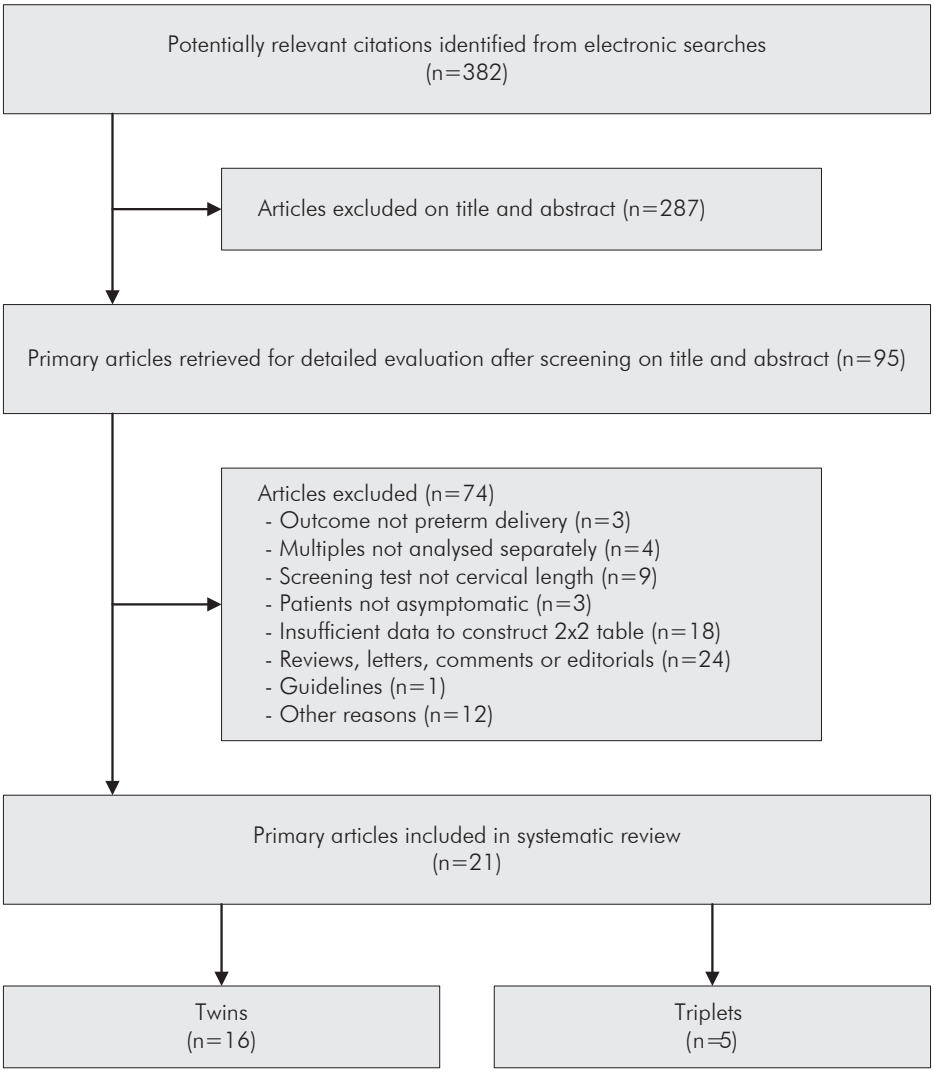

Figure 1 Flow chart to illustrate study method.

Quality assessment

Table 2 summarizes the results of the quality assessment. Of the 16 studies with an adequate description of the test, the study by Yang et al.[9] was the only one that included funneling in cervical length. Several authors referred to previous publications by others for a description of the testing technique. Most often cited in this context were studies by Andersen et al.[10] and Iams et al[11]. Practitioners were blinded to the results of cervical length measurement and no interventions based on cervical length were performed in only five studies. In 15 studies patients who had either a cerclage or an indicated preterm birth were excluded from the analysis.

Table 1 Study characteristics

Study	Inclusion criteria	Exclusion criteria	n	Gestational age at testing (weeks+days)	Cutoff points for cervical length (mm)	Cutoff points for gestational age at birth (weeks)
Asymptomatic twin pregnancies						
Arabin, 2006 [18]	twin pregnancy	iatrogenic PTB	153	15 - 19+6 20 - 24+6 25 - 29+6	15, 20, 25, 30	36
Fait, 2005 [19]	twin pregnancy selective reduction from triplet	fetal anomalies placental abruption Intrauterine growth retardation pre-eclampsia	20	15 - 17	35	33
Gibson, 2004 [20]	twin pregnancy	fetal anomalies TTTS	82	18 24 28 32	20, 22, 25	35
Goldenberg, 1996 [21]	twin pregnancy	cerclage fetal anomalies placenta praevia	147	22 - 24 27 - 28	25	32, 35, 37
Guzman, 20001 [22]	twin pregnancy	cerclage iatrogenic PTB	117	15 - 20 21 - 24 25 - 28	20	28, 30, 32, 34
Imseis, 1997 [23]	twin pregnancy	cerclage iatrogenic PTB	85	24 - 26	35	34
Klein, 2008 [24]	twin pregnancy	fetal anomalies ruptured membranes vaginal blood loss maternal pathology	223	20-25	25,30,35	34
McMahon, 2002 [25]	twin/triplet pregnancy	cerclage ruptured membranes iatrogenic PTB	109	20 24	30	32
Naba, 2000 [26]	twin pregnancy	-	14	24 - 34	25, 30	37
Robyr, 2005 [27]	twin pregnancy treated for TTTS	iatrogenic PTB	137	16 - 26	20, 25, 30, 35	28, 32, 34
Sayin, 2005 [28]	twin pregnancy	cerclage iatrogenic PTB	434	22 - 24	15, 20, 25	33
Soriano, 2002 [29]	twin pregnancy nulliparous ART	uterine anomalies DES exposure selective reduction iatrogenic PTB	44	18 - 24	35	34
Souka, 1999 [30]	twin pregnancy	cerclage TTTS	212	22 - 24	15, 25, 35, 45	28, 30, 32, 34
Sperling, 2005 [31]	twin pregnancy	iatrogenic PTB cerclage prior conization	383	23	21, 26, 31, 36	28, 32, 33, 34, 35
Venditelli, 2001 [32]	twin pregnancy	iatrogenic PTB	26	18 - 36	30	37

Chapter 7

Table 1 (*Continued*)

Study	Inclusion criteria	Exclusion criteria	n	Gestational age at testing (weeks+days)	Cutoff points for cervical length (mm)	Cutoff points for gestational age at birth (weeks)
Yang, 2000 [9]	twin pregnancy	cerclage placenta praevia preterm labor vaginal blood loss iatrogenic PTB	65	18 22 26	25, 30, 35	32, 35, 37
Asymptomatic triplet pregnancies						
Guzman, 20002 [33]	triplet pregnancy	cerclage iatrogenic PTB	50	15 - 20 21 - 24 25 - 28	20, 25	28, 30, 32
Maslovitz, 2004 [34]	triplet pregnancy trichorionic	selective reduction iatrogenic PTB	36	14 - 20	25	32
Maymon, 2001 [35]	triplet pregnancy	cerclage selective reduction uterine contractions	34	23 26 29	25	33
Missfelder-Lobos, 2003 [36]	triplet pregnancy	-	29	20 - 26	none	none
To, 2000 [37]	triplet pregnancy	-	38	22 - 24	15, 25, 30	33

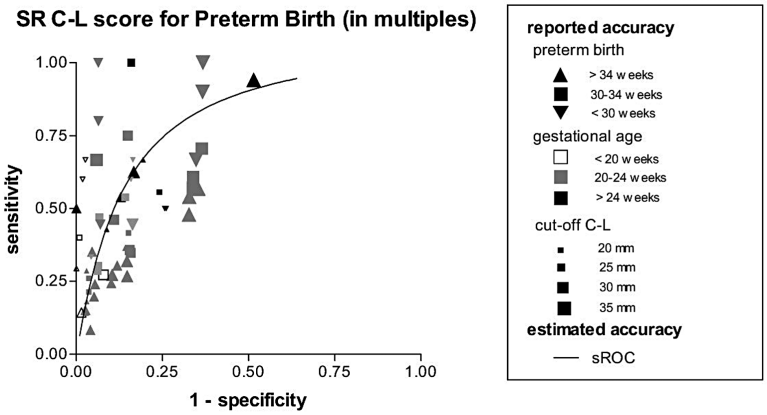

Figure 2 Summary receiver–operating characteristics curve for all studies (twin and triplet).

Data analysis

Figure 2 shows an ROC space of the individual studies, as well as the bootstrapped sROC curve plot for all studies in asymptomatic women with a twin or triplet pregnancy. Studies in which cervical length was measured before 20 weeks show an overall high specificity, but sensitivity in these studies does not exceed 70%. The overall sROC curve shows a moderate accuracy, with sensitivity being slightly better than specificity.

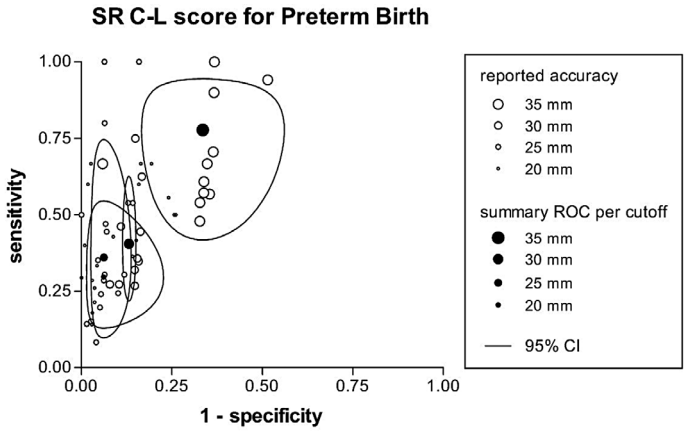

Figure 3 Summary receiver–operating characteristics space for different cut-off points of cervical length.

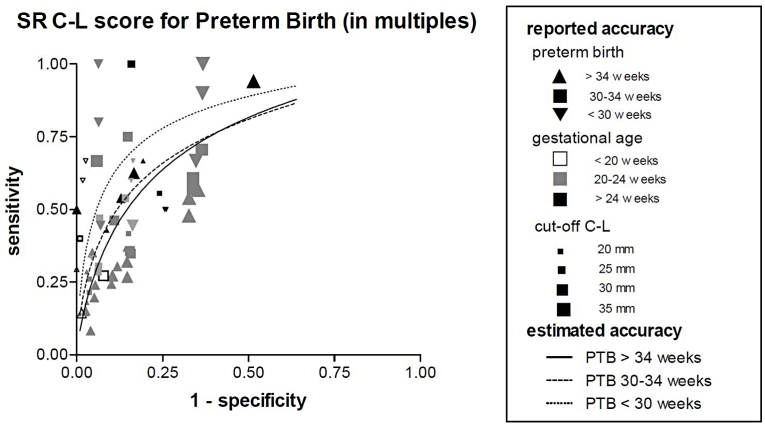

Figure 4 Summary receiver–operating characteristics curves for different cut-off points of preterm birth.

Figure 3 shows summary point estimates for sensitivity and specificity with 95% CIs for four different cut-offs of cervical length. The sensitivity and specificity for birth before 34 weeks are 78% and 66% for 35 mm, 41% and 87% for 30 mm, 36% and 94% for 25 mm and 30% and 94% for 20 mm. As is to be expected, cervical length cutoff is positively correlated with sensitivity, but negatively correlated with specificity.

Figure 4 shows the accuracy of the test per cut-off point for preterm birth. Accuracy was slightly better for preterm birth before 30 weeks than after 30 weeks and overall accuracy of cervical length measured after 24 weeks was slightly better than measurements before 24 weeks (results not shown).

Subgroup analysis showed that estimates of sensitivity and specificity were similar for twins (sensitivity 55%, specificity 90%) and triplets (sensitivity 51%, specificity 89%). When analysis was restricted to blinded studies only, cervical length measurement was less sensitive (35%), but this was not statistically significant. Exclusion from the analysis

Table 2 Study quality.

Study	Design	Adequate description of test procedure	Cut-off points determined with ROC-curve
Arabin, 2006 [18]	prospective cohort	yes	unclear
Fait, 2005 [19]	retrospective cohort	yes	no
Gibson, 2004 [20]	prospective cohort	yes	yes
Goldenberg, 1996 [21]	prospective cohort	yes	no
Guzman, 20001 [22]	prospective cohort	no	yes
Guzman, 20002 [33]	prospective cohort	no	yes
Imseis, 1997 [23]	retrospective cohort	yes	yes
Klein, 2008 [24]	retrospective cohort	yes	unclear
Maslovitz, 2004 [34]	retrospective cohort	yes	no
Maymon, 2001 [35]	prospective cohort	yes	yes
McMahon, 2002 [25]	prospective cohort	no	yes*
Missfelder-Lobos, 2003 [36]	retrospective cohort	yes	no†
Naba, 2000 [26]	prospective cohort	no	unclear
Robyr, 2005 [27]	retrospective cohort	no	no
Sayin, 2005 [28]	retrospective cohort	yes	no
Soriano, 2002 [29]	prospective cohort	yes	unclear
Souka, 1999 [30]	retrospective cohort	yes	no
Sperling, 2005 [31]	prospective cohort	yes	no
To, 2000 [37]	retrospective cohort	yes	no
Venditelli, 2001 [32]	prospective cohort	yes	no
Yang, 2000 [9]	prospective cohort	yes	no

* Cut-off point determined with ROC-curve at 27 mm, but adjusted to 30 mm for optimal clinical usefulness; † No cut-off points used, original data available ; ‡ Measurements ≥15 mm and higher were blinded to the caregiver, patients with measurements <15 mm received treatment

of the studies in which treatment for twin-to-twin transfusion syndrome (Robyr et al.[12]) or selective reduction (Fait et al.[13]) was performed, did not change estimates of sensitivity or specificity.

Discussion

We reviewed the literature on cervical length measurement for the prediction of preterm birth in asymptomatic women with a multiple pregnancy. Bivariate meta-analysis of the available data showed a strong association of short cervical length with preterm birth in multiple pregnancies. However, owing to a large variation in gestational age at measurement, cut-off points for cervical length and definitions of preterm birth, and due to the lack of effective treatments, no specific recommendations can be made for the application of this test in clinical practice. Sensitivity of the test is generally low and

Practitioner blinded to measurement results	Intervention based on cervical length	Exclusion of patients with intervention or indicated PTB
no	yes	yes
unclear	no	yes
yes	no	yes
yes	no	no
no	yes	yes
no	yes	yes
no	yes	yes
unclear	no	unclear
no	unclear	yes
unclear	unclear	yes
yes	no	yes
no	yes	unclear
unclear	unclear	no
no	yes	yes
no	yes	unclear
yes	no	yes
no	yes	yes
no‡	yes	yes
no	yes	yes
yes	no	no
no	yes	yes

increases with higher cut-off points for cervical length and preterm birth, at the expense of specificity. Specificity on the other hand is relatively high and increases with an earlier gestational age at testing.

This review has some strengths and limitations. We carried out extensive literature searches without language restrictions and systematically assessed the quality of studies. We have included all available studies reporting on the accuracy of cervical length, for a variety of definitions of preterm birth and thresholds for a positive test result and for different gestational ages at which cervical length was measured, as at present it is unclear which is the most appropriate value to use. This heterogeneity in study characteristics could not, however, be incorporated into the bivariate model commonly used for meta-analyses of diagnostic studies of test accuracy. An adjusted approach based on stratified bootstrap analysis was used to avoid the results being biased towards studies reporting accuracy for multiple thresholds for cervical length, preterm birth and/or gestational age at measurement. Until more sophisticated methods become available to deal with this within-study heterogeneity, this approach may be a viable alternative for evaluating the impact of this heterogeneity on diagnostic accuracy.

Many of these methodological problems may be overcome by using individual patient data meta analysis. From such data sets, population and patient characteristics, testing conditions and outcome variables can be made as uniform as possible before applying meta-analysis. Currently, initiatives in this field are being employed.[14] Individual patient data meta-analysis may also allow mono- and multichorionic twins to be analyzed separately, as the risk for preterm birth is strongly influenced by chorionicity.[15]

In most of the studies included in this review the results of cervical length measurement were not blinded to the caregiver, and interventions were performed based on these findings. Irrespective of whether or not patients with an intervention are included in the analysis, not blinding the caregiver causes information bias. However, we did not observe clinically relevant differences between the potentially biased studies and the studies in which information on cervical length measurement was unknown to the clinician managing the patient.

In spite of the large amount of research that has already been conducted in this subject area, effective strategies for the prevention of preterm birth in multiple pregnancies have yet to be established. Although in one trial progesterone appeared to be effective in both singleton and twin pregnancies with second-trimester cervical shortening, the proportion of twin pregnancies in that trial was only small (24 twin vs. 226 singleton pregnancies) and the effect of progesterone in women with a twin pregnancy was not statistically significant2. Three large trials in multiple pregnancies did not show a reduction in preterm birth after the use of progesterone.[5, 6, 16] In fact, there was a statistically non-significant trend towards more preterm births in the progesterone group in these trials. Cervical length measurements were not reported in any of the studies.

Two treatment strategies that have been explored in the past even seem to have a detrimental effect on unselected multiple pregnancies. A Cochrane review showed that women with an uncomplicated twin pregnancy who were hospitalized for bed rest had a significantly higher risk of delivering before 34 weeks (odds ratio (OR), 1.8; 95% CI, 1.01–3.3).[17] In an individual patient data meta-analysis, cerclage in multiple gestations was also found to increase pregnancy loss or death before discharge from the hospital (OR, 5.9; 95% CI, 1.1–30).[18] Another meta-analysis showed a significant increase in preterm birth at less than 35 weeks in twin gestations after cerclage (relative risk, 2.2; 95% CI, 1.2–4.0).[19]

An intervention that deserves further research is the vaginal pessary. Retrospective studies on this treatment have shown promising results, and several randomized controlled trials are currently being conducted.[20] One of these trials will include only women with a multiple pregnancy and will incorporate a cervical length measurement before the start of treatment in the second trimester.[21]

Despite the lack of an effective intervention seen in unselected populations of women with twin pregnancy, the strong predictive value of cervical length measurement for preterm birth possibly allows the identification of a subgroup of women with a multiple

pregnancy that can benefit from preventive treatments such as progesterone. Past efforts to find a preventive treatment for preterm birth have mostly focused on women with a multiple pregnancy in general. Although it is true that this entire group is at increased risk for preterm birth as compared to singleton pregnancies, it is very likely that the individual risk varies and that a subgroup of women with a multiple pregnancy is at an even higher risk for delivering preterm. Just as progesterone and cerclage only seem to have an effect in selected subgroups of women with a singleton pregnancy, these and other interventions may only be effective in a selected proportion of women with a multiple pregnancy.

In this meta-analysis we did not address the risk of delivery within 14 days based on cervical length. This could be an interesting subject for future research, as cervical length may help to select patients who are likely to benefit from treatment with a course of corticosteroids.

In summary, the results of this meta-analysis show that in asymptomatic women with a multiple pregnancy, measurement of second-trimester cervical length can be used to identify a group of women who are at increased risk for preterm birth. Sensitivity, however, is low, indicating that a large percentage of women with a multiple pregnancy will deliver prematurely in spite of a long cervix in the second trimester. In view of the fact that cerclage is known to increase complications, a high specificity is also important, as a limited specificity would increase the exposure of women with a low risk of preterm delivery to unnecessary and potentially harmful interventions. For example, a cut-off of 35 mm with a sensitivity of 78% and specificity of 66%, and a prevalence of preterm delivery of 20% would indicate that 60% of women with a cervix shorter than 35 mm would not deliver before 34 weeks. We think that in future trials on preventive strategies for preterm birth in multiple pregnancies, a blinded second trimester measurement of cervical length should be part of the study protocol. In this way, planned subgroup analyses can ascertain whether women with a short cervix will benefit from treatment more than others. If preventive treatments are then found to be effective in these women, cervical length measurement could establish a role in the clinical management of women with a multiple pregnancy.

References

(1) 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(3):305-322.

(2) Fonseca EB, Celik E, Parra M, Singh M, Nicolaides KH. Progesterone and the risk of preterm birth among women with a short cervix. *N Engl J Med* 2007;357(5):462-469.

(3) da Fonseca EB, Bittar RE, Carvalho MH, Zugaib M. Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: a randomized placebo-controlled double-blind study. *Am J Obstet Gynecol* 2003;188(2):419-424.

(4) Meis PJ, Klebanoff M, Thom E et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. *N Engl J Med* 2003;348(24):2379-2385.

(5) Norman JE, Mackenzie F, Owen P et al. Progesterone for the prevention of preterm birth in twin pregnancy (STOPPIT): a randomised, double-blind, placebo-controlled study and meta-analysis. *Lancet* 2009;373(9680):2034-2040.

(6) Rouse DJ, Caritis SN, Peaceman AM et al. A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. *N Engl J Med* 2007;357(5):454-461.

(7) 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.

(8) Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58(10):982-990.

(9) Yang JH, Kuhlman K, Daly S, Berghella V. Prediction of preterm birth by second trimester cervical sonography in twin pregnancies. *Ultrasound Obstet Gynecol* 2000;15(4):288-291.

(10) Andersen HF, Nugent CE, Wanty SD, Hayashi RH. Prediction of risk for preterm delivery by ultrasonographic measurement of cervical length. *Am J Obstet Gynecol* 1990;163(3):859-867.

(11) 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(9):567-572.

(12) Robyr R, Boulvain M, Lewi L et al. Cervical length as a prognostic factor for preterm delivery in twin-to-twin transfusion syndrome treated by fetoscopic laser coagulation of chorionic plate anastomoses. *Ultrasound Obstet Gynecol* 2005;25(1):37-41.

(13) Fait G, Har-Toov J, Gull I, Lessing JB, Jaffa A, Wolman I. Cervical length, multifetal pregnancy reduction, and prediction of preterm birth. *J Clin Ultrasound* 2005;33(7):329-332.

(14) Broeze KA, Opmeer BC, Bachmann LM et al. Individual patient data meta-analysis of diagnostic and prognostic studies in obstetrics, gynaecology and reproductive medicine. *BMC Med Res Methodol* 2009;9:22.

(15) Hack KE, Derks JB, de Visser VL, Elias SG, Visser GH. The natural course of monochorionic and dichorionic twin pregnancies: a historical cohort. *Twin Res Hum Genet* 2006;9(3):450-455.

(16) Caritis SN, Rouse DJ, Peaceman AM et al. Prevention of preterm birth in triplets using 17 alpha-hydroxyprogesterone caproate: a randomized controlled trial. *Obstet Gynecol* 2009;113(2 Pt 1):285-292.

(17) Crowther CA, Han S. Hospitalisation and bed rest for multiple pregnancy. *Cochrane Database Syst Rev* 2010;(7):CD000110.

(18) Jorgensen AL, Alfirevic Z, Tudur SC, Williamson PR. Cervical stitch (cerclage) for preventing pregnancy loss: individual patient data meta-analysis. *BJOG* 2007;114(12):1460-1476.

(19) Berghella V, Odibo AO, To MS, Rust OA, Althuisius SM. Cerclage for short cervix on ultrasonography: meta-analysis of trials using individual patient-level data. *Obstet Gynecol* 2005;106(1):181-189.

(20) Dharan VB, Ludmir J. Alternative treatment for a short cervix: the cervical pessary. *Semin Perinatol* 2009;33(5):338-342.

(21) Hegeman MA, Bekedam DJ, Bloemenkamp KW et al. Pessaries in multiple pregnancy as a prevention of preterm birth: the ProTwin Trial. *BMC Pregnancy Childbirth* 2009;9:44.