






Prognostic accuracy of cerebroplacental ratio and middle cerebral artery Doppler for adverse perinatal outcome: systematic review and meta-analysis

C. A. VOLLGRAFF HEIDWEILLER-SCHREURS¹ , M. A. DE BOER¹, M. W. HEYMANS²,
L. J. SCHOONMADE³ , P. M. M. BOSSUYT⁴, B. W. J. MOL^{5,6} , C. J. M. DE GROOT¹
and C. J. BAX⁷

¹Department of Obstetrics and Gynecology, VU University Medical Center, Amsterdam, The Netherlands; ²Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands; ³Department of Medical Library, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; ⁴Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; ⁵Department of Obstetrics and Gynecology, The Robinson Institute, School of Medicine, University of Adelaide, Adelaide, Australia; ⁶The South Australian Health and Medical Research Institute, Adelaide, Australia; ⁷Department of Obstetrics and Gynecology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

KEYWORDS: cerebroplacental ratio; Doppler; fetal growth restriction; middle cerebral artery; prognostic accuracy

ABSTRACT

Objective Doppler ultrasonographic assessment of the cerebroplacental ratio (CPR) and middle cerebral artery (MCA) is widely used as an adjunct to umbilical artery (UA) Doppler to identify fetuses at risk of adverse perinatal outcome. However, reported estimates of its accuracy vary considerably. The aim of this study was to review systematically the prognostic accuracies of CPR and MCA Doppler in predicting adverse perinatal outcome, and to compare these with UA Doppler, in order to identify whether CPR and MCA Doppler evaluation are of added value to UA Doppler.

Methods PubMed, EMBASE, the Cochrane Library and ClinicalTrials.gov were searched, from inception to June 2016, for studies on the prognostic accuracy of UA Doppler compared with CPR and/or MCA Doppler in the prediction of adverse perinatal outcome in women with a singleton pregnancy of any risk profile. Risk of bias and concerns about applicability were assessed using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) tool. Meta-analysis was performed for multiple adverse perinatal outcomes. Using hierarchical summary receiver–operating characteristics meta-regression models, the prognostic accuracy of CPR vs MCA Doppler was compared indirectly, and CPR and MCA Doppler vs UA Doppler compared directly.

Results The search identified 4693 articles, of which 128 studies (involving 47 748 women) were included. Risk of bias or suboptimal reporting was detected in 120/128 studies (94%) and substantial heterogeneity was found, which limited subgroup analyses for fetal growth and gestational age. A large variation was observed in reported sensitivities and specificities, and in thresholds used. CPR outperformed UA Doppler in the prediction of composite adverse outcome (as defined in the included studies) ($P < 0.001$) and emergency delivery for fetal distress ($P = 0.003$), but was comparable to UA Doppler for the other outcomes. MCA Doppler performed significantly worse than did UA Doppler in the prediction of low Apgar score ($P = 0.017$) and emergency delivery for fetal distress ($P = 0.034$). CPR outperformed MCA Doppler in the prediction of composite adverse outcome ($P < 0.001$) and emergency delivery for fetal distress ($P = 0.013$).

Conclusion Calculating the CPR with MCA Doppler can add value to UA Doppler assessment in the prediction of adverse perinatal outcome in women with a singleton pregnancy. However, it is unclear to which subgroup of pregnant women this applies. The effectiveness of the CPR in guiding clinical management needs to be evaluated in clinical trials. © 2017 The Authors. *Ultrasound in Obstetrics & Gynecology* published by John Wiley & Sons Ltd on behalf of the International Society of *Ultrasound in Obstetrics and Gynecology*.

Correspondence to: Dr C. A. Vollgraaf Heidweiller-Schreurs, Department of Obstetrics and Gynecology, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands (e-mail: c.schreurs@vumc.nl)

Accepted: 7 July 2017

INTRODUCTION

In fetal growth restriction (FGR), Doppler ultrasound examination can be used to distinguish between fetuses that are at risk of adverse perinatal outcome and those that are constitutionally small. Current FGR guidelines recommend umbilical artery (UA) Doppler as an important surveillance tool^{1–3}, since its clinical effectiveness in high-risk pregnancies has been reported in a Cochrane review of 18 randomized controlled trials⁴.

Fetal middle cerebral artery (MCA) Doppler has been proposed as an additional test to UA Doppler. Decreased impedance in the MCA is considered the ‘brain-sparing’ effect; a manifestation of redistribution of the fetal circulation and possibly a sign of compromise. A systematic review of MCA Doppler reported limited summary likelihood ratios for predicting adverse perinatal outcome⁵. The cerebroplacental ratio (CPR) is calculated as the ratio of MCA to UA Doppler, and has been hypothesized to be more accurate than its individual components⁶. An association with adverse perinatal outcome has been the focus of several literature reviews^{7–10}. CPR and MCA Doppler in FGR are gradually becoming integrated into clinical practice and international guidelines^{1,3,11}, but reported estimates of their accuracy vary considerably.

We conducted a systematic review and meta-analysis of published studies on the prognostic accuracy of CPR and MCA Doppler compared with UA Doppler in the prediction of adverse perinatal outcome, in order to identify whether CPR and MCA Doppler evaluation are of added value to UA Doppler.

METHODS

Search strategy and selection criteria

This systematic review is described in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (www.prisma-statement.org). The protocol for this systematic review was registered in the PROSPERO database (CRD42016039915).

A comprehensive search of PubMed, EMBASE, the Cochrane Library (via Wiley) and ClinicalTrials.gov was performed from inception to June 14th 2016, in collaboration with a medical librarian. Search terms included controlled terms (MeSH in PubMed, Emtree in EMBASE) as well as free-text terms (Cochrane Library only). The following terms were used as index terms or free-text words (including synonyms and closely related words): (‘pregnancy’ or ‘fetus’) and (‘middle cerebral artery’ or ‘CPR’ or ‘brain sparing’). The full search strategies for all the databases can be found in Appendix S1. Duplicate articles were excluded and reference lists of included articles and previous reviews scanned for further eligible articles. Articles were not filtered based on language.

We searched for studies that assessed the prognostic accuracy of UA Doppler compared with CPR and/or MCA Doppler in women with a singleton pregnancy of any risk profile, without known chromosomal or

structural abnormalities. Studies were eligible if they reported on the association between Doppler indices and at least one of the following adverse perinatal outcomes: perinatal death, 5-min Apgar score < 7, acidosis on cord blood gas analysis, abnormal cardiocotogram, emergency delivery for fetal distress, meconium-stained fluid, need for assisted respiration after birth, need for resuscitation, admission to the neonatal intensive care unit (NICU) or neonatal unit, serious neonatal morbidity (e.g. necrotizing enterocolitis, sepsis, respiratory distress syndrome), neurological outcomes, ultrasound-detected neonatal intracranial abnormalities, a composite of adverse outcomes (as defined in the included studies) or birth weight (BW) < 10th percentile. In cases of multiple publications from one study group, we included the data reporting on the largest study group. Studies reporting on UA Doppler but not on MCA Doppler or CPR were excluded, as were review articles and case reports.

Data extraction and quality assessment

Two reviewers (C.A.V. and M.A.B.) independently assessed titles and abstracts of all search results. Full reports of studies considered potentially eligible by at least one of the reviewers were obtained. The same two reviewers (C.A.V. and M.A.B.) assessed the full reports for inclusion and extracted data. In case of doubt, other reviewers were consulted (C.J.B., B.W.J.M. or C.J.M.G.).

From each included study report, we identified the first author, country, journal, year of publication, recruitment setting, sample size and characteristics of included patients (age, gestational age at test and at delivery, fetal growth and gestational hypertensive disorders). We extracted data on the index test (including Doppler index, and threshold and reference values), targeted outcomes, accuracy estimates and data for 2 × 2 tables. Whenever UA Doppler results were presented, we also extracted data for 2 × 2 tables. Risk of bias and concerns about applicability were assessed by two authors (C.A.V. and C.J.B.) with the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) tool¹².

Statistical analysis

Meta-analyses were performed for the outcomes perinatal death, 5-min Apgar score < 7, emergency delivery for fetal distress, NICU admission and composite adverse perinatal outcome (as defined in the included studies). These were selected by consensus, as they are the most homogeneously classified and least subject to ascertainment bias. Owing to variation in thresholds across studies, the hierarchical summary receiver–operating characteristics (HSROC) model of Rutter and Gatsonis^{13,14} was used to obtain summary ROC curves of the CPR and MCA Doppler for studies that had evaluated the same target outcome. The HSROC model allows for variation of accuracy and thresholds between studies. One threshold per study was selected to be included in this analysis. If a study reported on more than one threshold, the most commonly

used Doppler index (i.e. pulsatility index (PI)) and threshold (i.e. for CPR < 1.08 and for MCA Doppler < 2 SD below the mean) were selected. Whenever three or more studies were available, estimates of mean sensitivity and specificity and respective variances at a specific threshold were additionally generated using the bivariate logit-normal model¹⁵.

The prognostic test accuracy of CPR *vs* MCA Doppler was compared indirectly, using a HSROC meta-regression model by including test type as a covariate. This meta-regression approach was also used to investigate heterogeneity, by adding potential sources of heterogeneity as covariates to the HSROC model. *A priori* considered covariates were: fetal growth (small-for-gestational age (SGA) and appropriate-for-gestational age (AGA)), duration of pregnancy (preterm and term) and timing of measurement (within 1 week of delivery and more than 1 week from delivery). Preterm delivery was defined as delivery before 37 weeks' gestation. SGA was defined as estimated fetal weight (EFW) or BW $< 10^{\text{th}}$ percentile.

The prognostic test accuracies of CPR and MCA Doppler were then compared directly with that of UA Doppler using a HSROC meta-regression model. For this purpose, analyses were restricted to studies that had compared two tests in the same study group, which ensures that differences in accuracy are not caused by heterogeneity in performance across study populations. Review Manager (RevMan) version 5.3 (Cochrane Collaboration, Copenhagen, 2014) was used for data extraction and generating forest plots and summary ROC curves. SAS® Studio (SAS Institute Inc., Cary, NC, USA) was used for HSROC model analysis and Stata version 14 (StataCorp LP, College Station, TX, USA) for the bivariate models.

RESULTS

Search results

A total of 4689 articles were identified through the electronic search, and four additional articles through cross-referencing (Figure 1). After removing duplicates, the remaining 2972 titles and abstracts were screened, resulting in 578 potentially eligible articles. Of these, 450 articles could not be included, for reasons listed in Figure 1. This resulted in 128 studies (involving 47 748 women) being included (Appendix S2), of which 52 studies evaluated CPR only, 48 studies MCA Doppler only and 28 studies both CPR and MCA Doppler.

Characteristics of included studies

Detailed characteristics of the included studies are provided in Table S1. Data had been prospectively collected in 73 studies (57%), retrospectively in 25 studies (20%), and by unclear methods in 30 studies (23%). The number of patients included in the different studies ranged from 19 to 9198. Mean or median maternal age ranged from 23.0 to 37.7 years. Definitions of FGR varied, ranging

from EFW $< 3^{\text{rd}}$ percentile with abnormal UA Doppler to BW $< 10^{\text{th}}$ percentile. For this review, we combined these definitions into SGA. Most studies investigated SGA fetuses (65 studies, 51%), while 15 studies (12%) investigated only AGA fetuses. In most studies, gestational age at examination was unknown or there was a range of gestational ages (78 studies, 61%), preterm pregnancies were investigated in 28 studies (22%) and term pregnancies were investigated in 22 (17%) studies.

Quality of included studies

Results of the QUADAS-2 assessment are provided in Appendix S3. Risk of bias or suboptimal reporting was detected in 120/128 studies (94%). The interval between index test and outcome was often larger than 1 week (53 studies (41%)) or unclear (37 studies (29%)). Methods for patient selection were also often unclear (36 studies (28%)). In 64 studies (50%) it was unclear whether the obstetrician was blinded to the test results, and in 30 studies (23%) they were not blinded; nine of these described if and how an abnormal test result influenced obstetric decision-making.

Meta-analysis and indirect comparisons of CPR and MCA Doppler

For the meta-analysis, we could include 106/128 studies assessing one of the five primary outcomes defined in the Methods section. In Figure 2, summary curves of all included studies estimated with the HSROC model are shown for CPR and MCA Doppler for the outcomes perinatal death, 5-min Apgar score < 7 , emergency delivery for fetal distress, NICU admission and composite adverse perinatal outcomes (as defined in the included studies). Forest plots of all other outcomes can be found in Figure S1. Clearly observable is the large variation in sensitivity and specificity. For perinatal death, the CPR was evaluated in 25 studies (298/8003 patients) and MCA Doppler in 28 studies (326/12 230 patients). For low Apgar score, CPR was evaluated in 28 studies (751/6904 patients) and MCA Doppler in 23 studies (329/12 889 patients). For emergency delivery for fetal distress, CPR was evaluated in 28 studies (3007/15 834 patients), and MCA Doppler in 26 studies (1745/14 171 patients). For NICU admission, CPR was evaluated in 28 studies (1632/5417 patients), and MCA Doppler in 20 studies (695/2224 patients). For composite adverse outcome, CPR was evaluated in 31 studies (1574/5271 patients), and MCA Doppler in 28 studies (1445/5477 patients). In these indirect comparisons, CPR performed better than did MCA Doppler for all outcomes (composite adverse outcome, $P < 0.001$; perinatal death, $P = 0.004$; emergency delivery for fetal distress, $P = 0.004$; and low Apgar score, $P = 0.033$), except for admission to the NICU ($P = 0.761$).

Thresholds

The asymmetric shape of the curves in Figure 2 suggests that accuracy varies with threshold. Many different

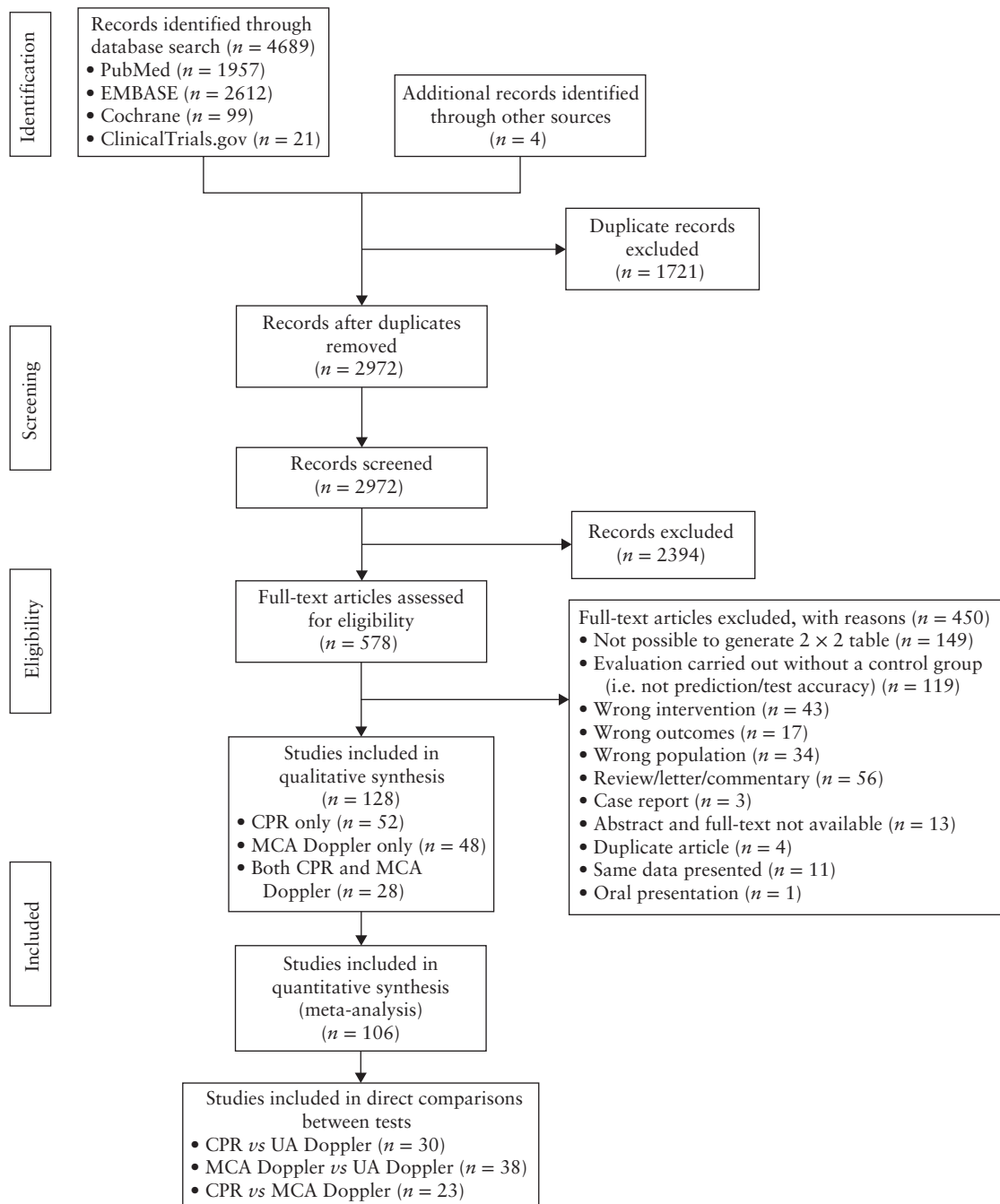


Figure 1 Flowchart of studies included in systematic review and meta-analysis on prognostic accuracy of cerebropoplental ratio (CPR) and/or middle cerebral artery (MCA) Doppler in prediction of adverse perinatal outcome in singleton pregnancies compared with umbilical artery (UA) Doppler.

thresholds were used in the studies, of which most were predefined (120 studies, 94%). Some studies used absolute thresholds, while others used percentiles on a reference curve (Appendix S4). Table 1 shows mean sensitivities and specificities, with their confidence and prediction intervals, of CPR and MCA Doppler at thresholds that were assessed in at least three studies for the outcomes perinatal death, 5-min Apgar score < 7, emergency delivery for fetal distress, NICU admission and composite adverse perinatal outcomes (as defined in the included studies). For this purpose, absolute thresholds were grouped from 1.0 to

1.1. The thresholds that could be taken together were CPR-PI 1.0–1.1, CPR resistance index 1.0–1.1, MCA-PI < 2 SD on the reference curve of Mari and Deter¹⁶, and MCA-PI < 5th percentile on the reference curve of Arduini and Rizzo¹⁷. Mean sensitivity of the CPR ranged from 0.55 to 0.93, and mean specificity ranged from 0.63 to 0.91. Mean sensitivity of MCA Doppler ranged from 0.34 to 0.67, and mean specificity ranged from 0.65 to 0.96. Prediction intervals were large, which indicates high between-study variability in sensitivity and specificity, and in thresholds.

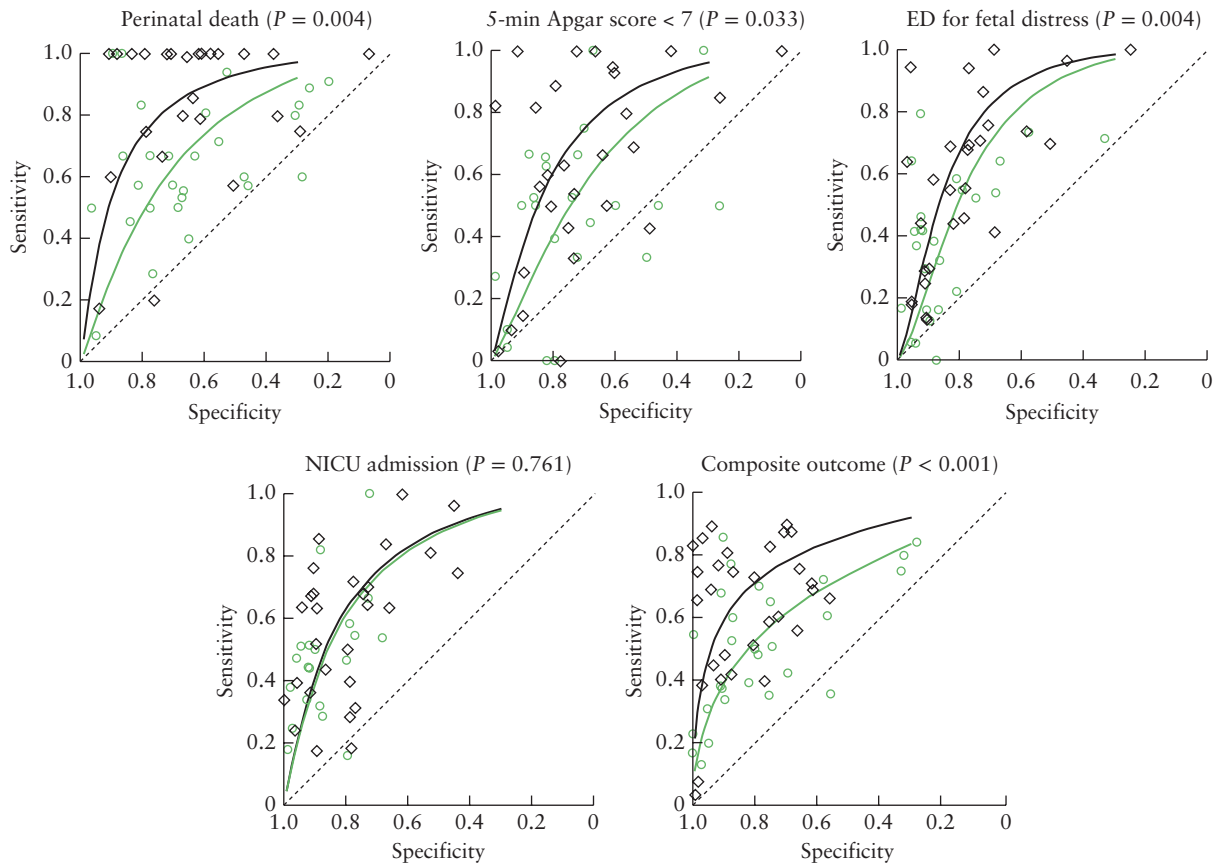


Figure 2 Hierarchical summary receiver–operating characteristics curves and *P*-values for indirect comparisons of prognostic accuracy of cerebroplacental ratio (◇, black line) and middle cerebral artery Doppler (○, green line) for outcomes perinatal death, 5-min Apgar score < 7, emergency delivery (ED) for fetal distress, admission to neonatal intensive care unit (NICU) and composite adverse perinatal outcome (as defined in included studies).

Table 1 Mean accuracy estimates of cerebroplacental ratio (CPR) and middle cerebral artery (MCA) Doppler for different adverse outcomes, based on fixed thresholds investigated in at least three studies

Index test; threshold	Outcome	Studies (n)	Patients (n)	Sensitivity			Specificity		
				Estimated mean	95% CI	95% predict int	Estimated mean	95% CI	95% predict int
CPR-PI; 1.0–1.1	Perinatal death	10	3571	0.93	0.71–0.99	0.15–0.99	0.74	0.62–0.84	0.32–0.95
	5-min Apgar score < 7	9	1370	0.61	0.53–0.68	0.47–0.73	0.76	0.69–0.82	0.55–0.89
	ED for fetal distress	8	2559	0.58	0.39–0.74	0.15–0.92	0.89	0.79–0.94	0.52–0.98
	NICU admission	12	2603	0.55	0.44–0.65	0.24–0.83	0.85	0.79–0.90	0.60–0.95
	Composite adverse outcome	13	2591	0.59	0.44–0.73	0.15–0.92	0.91	0.82–0.96	0.39–0.99
CPR-RI; 1.0–1.1	Perinatal death	5	450	0.84	0.63–0.94	0.45–0.97	0.67	0.45–0.84	0.19–0.95
	5-min Apgar score < 7	8	1782	0.75	0.42–0.93	0.06–0.99	0.84	0.63–0.94	0.16–0.99
	NICU admission	4	299	0.66	0.46–0.82	0.30–0.76	0.63	0.49–0.76	0.36–0.84
	Composite adverse outcome	4	409	0.80	0.69–0.89	0.59–0.93	0.85	0.66–0.94	0.40–0.98
MCA-PI; < 2 SD ¹⁶	Perinatal death	3	92	0.48	0.28–0.68	0.27–0.69	0.72	0.53–0.85	0.40–0.91
	5-min Apgar score < 7	3	125	0.56	0.35–0.75	0.32–0.77	0.65	0.45–0.80	0.31–0.88
	ED for fetal distress	3	167	0.52	0.36–0.68	0.28–0.75	0.84	0.51–0.97	0.25–0.99
	NICU admission	5	385	0.45	0.33–0.57	0.23–0.69	0.93	0.84–0.97	0.70–0.99
	Composite adverse outcome	3	211	0.34	0.18–0.55	0.10–0.69	0.96	0.77–0.99	0.45–1.00
MCA-PI; < 5 th per- centile ¹⁷	ED for fetal distress	3	416	0.35	0.22–0.49	0.16–0.59	0.83	0.78–0.87	0.77–0.87
	NICU admission	3	270	0.64	0.10–0.99	0.02–0.99	0.75	0.67–0.81	0.51–0.89
	Composite adverse outcome	3	511	0.67	0.58–0.74	0.55–0.75	0.78	0.59–0.89	0.39–0.95

ED, emergency delivery; NICU, neonatal intensive care unit; PI, pulsatility index; predict int, prediction interval; RI, resistance index.

Heterogeneity between subgroups

We found evidence of heterogeneity (Table S2) in the overall test performance of the CPR related to fetal growth (SGA vs AGA), with only one out of five outcomes showing a statistically significant different accuracy (in favor of AGA). The forest plots (Figure S1) show that, in general, sensitivity is lower and specificity is similar in AGA studies compared with SGA studies. We also found evidence of heterogeneity (Table S2) in the overall test performance of the CPR and MCA Doppler related to time of pregnancy (preterm vs at term), with only two out of five outcomes showing statistically significantly different accuracy (in favor of preterm pregnancies). We found no statistical evidence of heterogeneity related to timing of the Doppler measurement (within or beyond 1 week of delivery). Other potential sources of heterogeneity, such as late FGR, could not be evaluated statistically because of the limited number of studies that reported on these covariates and owing to heterogeneity in outcome reporting. In late SGA (≥ 32 weeks' gestation), nine studies investigated the prognostic accuracy of the CPR and eight investigated the prognostic accuracy of MCA Doppler for different outcomes. These studies are summarized in Appendix S5. For the CPR, reported sensitivity ranged from 0.08 to 0.71, while specificity ranged from 0.47 to 0.98. For MCA Doppler, reported sensitivity ranged from 0.13 to 1.00, while specificity ranged from 0.67 to 0.97.

Direct comparisons of CPR, MCA Doppler and UA Doppler

Figure 3 shows direct comparisons of the prognostic accuracy of CPR, MCA Doppler and UA Doppler for the outcomes perinatal death, 5-min Apgar score < 7 , emergency delivery for fetal distress, NICU admission and composite adverse perinatal outcomes (as defined in the included studies). Overall, the CPR and UA Doppler were compared in 30 studies (5046 patients), MCA Doppler and UA Doppler were compared in 38 studies (18 999 patients), and the CPR and MCA Doppler were compared in 23 studies (4262 patients). The connecting lines between the pair of tests from each study appear to show that sensitivity of CPR outperforms that of UA Doppler and sensitivity of MCA Doppler performs less well than do those of UA Doppler and CPR. A meta-regression analysis was performed to compare statistically prognostic accuracy of the three tests in the studies that had performed a direct comparison. The CPR was significantly better than UA Doppler in predicting composite adverse outcome ($P < 0.001$) and emergency delivery for fetal distress ($P = 0.003$), but comparable with UA Doppler in predicting perinatal death ($P = 0.686$), low Apgar score ($P = 0.595$) and NICU admission ($P = 0.107$). MCA Doppler was significantly worse than UA Doppler in predicting low Apgar score ($P = 0.017$) and emergency delivery for fetal distress ($P = 0.034$) and significantly worse than CPR in

predicting composite adverse outcome ($P < 0.001$) and emergency delivery for fetal distress ($P = 0.013$).

DISCUSSION

Summary of main results

We reviewed systematically studies on the prognostic accuracy of CPR and/or MCA Doppler compared with UA Doppler in predicting adverse perinatal outcome, to ascertain whether CPR and MCA Doppler evaluation are of added value to UA Doppler. Formal quality assessment of the included studies revealed few studies of high quality. We observed a large variation in thresholds used and in reported sensitivities and specificities. In the direct test comparisons, prognostic accuracy of the CPR significantly outperformed that of UA Doppler for emergency delivery for fetal distress and composite adverse outcome, while for the other outcomes CPR was similar to UA Doppler. For all outcomes, sensitivity of CPR appeared to be better than that of UA Doppler. Prognostic accuracy of MCA Doppler was significantly worse than those of UA Doppler and CPR for most outcomes.

Strengths and limitations of the study

The main strength of this review is that its conclusions are based on direct within-study comparisons in a large dataset, which limits confounding. However, several limitations deserve consideration. Subgroup analyses were limited owing to the small number of studies that explicitly described inclusion criteria and heterogeneity in outcome reporting. Almost all included studies were at risk of bias and had suboptimal reporting. Caregivers in almost all studies were aware of the Doppler results, which may have influenced obstetric decision-making. Another concern is that publication bias may have led to overestimation of accuracy. Although no appropriate method exists to quantify publication bias in test accuracy reviews¹⁸, studies with higher sensitivities and specificities presumably have better chances of being published than those with lower sensitivities and specificities. Furthermore, the optimal test, threshold and outcome may have been reported selectively.

The large differences in test accuracy between studies indicate substantial heterogeneity, for which several sources should be considered. Study populations varied, ranging from early severe FGR to uncomplicated term pregnancies. The frequency of adverse perinatal outcomes also fluctuated. Another source of heterogeneity is the large variety of thresholds used for test positivity across studies. Several methods were used across studies to classify adverse perinatal outcome and we performed a meta-analysis only for those that could be classified homogeneously, aiming to minimize this source of heterogeneity.

Implications for practice

In current practice, the management of FGR is aimed at monitoring fetal condition in order to time optimally

induction of delivery. In early FGR (< 32 weeks' gestation), clinical management is guided primarily by Doppler measurements of the UA and ductus venosus¹. In late FGR (≥ 32 weeks), it has been suggested that CPR and MCA Doppler may be of specific clinical value^{8–11}.

Late-onset FGR is caused by a milder degree of placental insufficiency and is more difficult to distinguish from the constitutionally small fetus. At present, there is a lack of evidence for the optimal timing of delivery in these pregnancies. A randomized clinical trial in women with

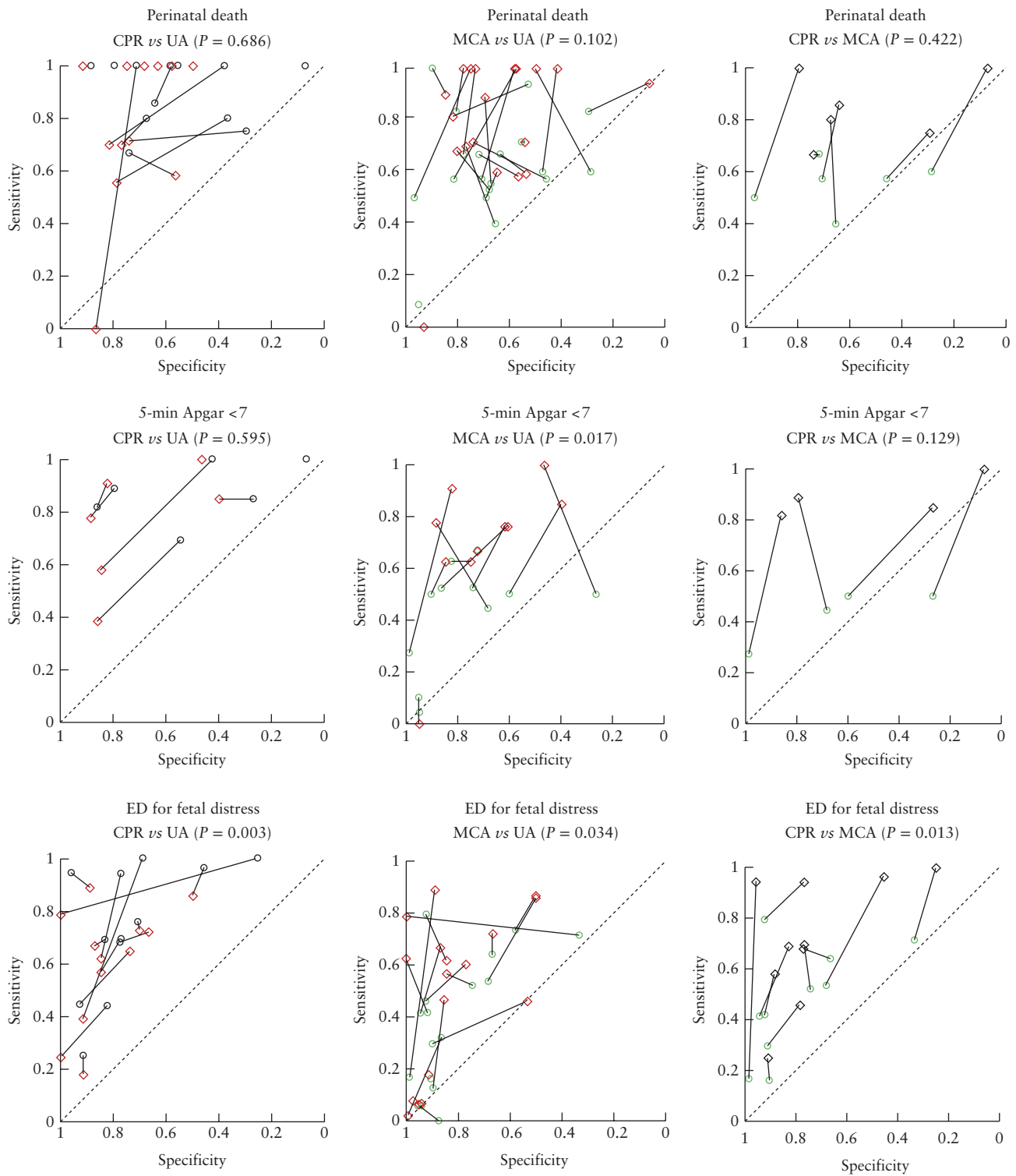


Figure 3 Hierarchical summary receiver–operating characteristics curves and *P*-values for direct comparisons of prognostic accuracy of cerebroplacental ratio (○), middle cerebral artery Doppler (○) and umbilical artery Doppler (◇) for outcomes perinatal death, 5-min Apgar score < 7, emergency delivery (ED) for fetal distress, admission to neonatal intensive care unit (NICU) and composite adverse perinatal outcome (as defined in included studies). Analyses restricted to studies that compared both tests in the same patients. Lines connect pairs of points representing the two tests from each study. (*Continued over.*)

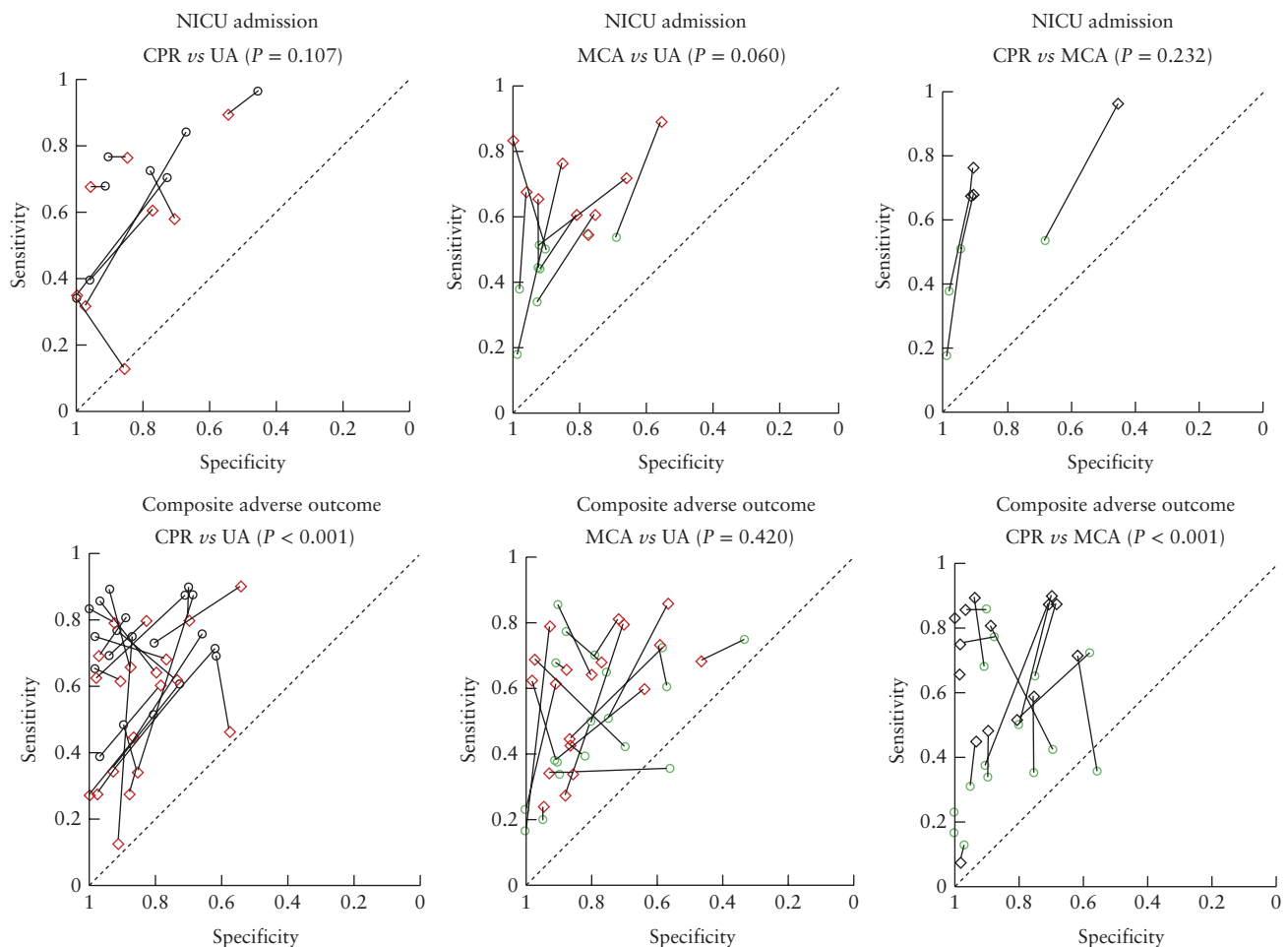


Figure 3 (Continued) Hierarchical summary receiver–operating characteristics curves and *P*-values for direct comparisons of prognostic accuracy of cerebraloplacental ratio (○), middle cerebral artery Doppler (○) and umbilical artery Doppler (◇) for outcomes perinatal death, 5-min Apgar score < 7, emergency delivery (ED) for fetal distress, admission to neonatal intensive care unit (NICU) and composite adverse perinatal outcome (as defined in included studies). Analyses restricted to studies that compared both tests in the same patients. Lines connect pairs of points representing the two tests from each study.

FGR at term indicated that induction of labor is not harmful¹⁹. A large national cohort study indicated that, for suspected FGR at 37 weeks, expectant management for a further week leads to less mortality than delivery, while continuation of pregnancy after 38 weeks is related to an increased risk, although this trend only reached statistical significance from 40 weeks onwards (relative risk, 2.46 (95% CI, 1.80–3.36))²⁰.

This review found that studies investigating CPR and MCA Doppler in late FGR are scarce, and that they evaluated different perinatal outcomes, which precluded a meta-analysis in this specific group. For pregnancies with EFW or BW < 10th percentile at ≥ 32 weeks, nine studies investigated prognostic accuracy of the CPR and eight that of MCA Doppler. We observed a large variation in accuracy of both the CPR and MCA Doppler and, with the available data, were not able to support the theory that the CPR or MCA Doppler is of particular clinical value in late FGR.

Another clinical role of CPR and MCA Doppler has been suggested in low-risk AGA pregnancy⁸. In this subgroup, we observed predominantly similar prognostic

accuracy to that in SGA pregnancies, although sensitivity appeared to decrease. Routinely evaluating CPR or MCA Doppler in a low-risk population with a low prevalence of adverse outcomes would lead to a substantial number of false negatives, and therefore would leave many at-risk fetuses undetected. Moreover, false positives could lead to unnecessary and potentially harmful interventions in this population. Unless the use of CPR or MCA Doppler is validated by clinical trials, we cannot recommend them in this setting.

Implications for research

The reproducibility and interpretation of observational studies on CPR could be improved in the future by using set inclusion criteria, completeness in reporting and by agreeing on standard outcome measures for FGR by consensus^{21,22}. We were unable to analyze the incremental value of combining UA Doppler with CPR, owing to a lack of studies specifically addressing this question. This could be answered by obtaining individual participant data. The logical next step is to evaluate

further the effectiveness of the CPR in clinical trials. Randomized clinical trials on a prognostic topic, such as sonographic markers, have their own challenges. The trial would be effective only if the clinician follows the result of the prognostic test under study, and the test would be beneficial only if subsequent interventions are effective²³. It is reasonable to use a higher threshold for CPR, for instance the 10th or even 20th percentile, in order to allow identification of the optimal threshold. Preferably, studies should measure both UA Doppler and CPR and randomly allocate women with discordant test results (i.e. normal UA Doppler and abnormal CPR, or vice versa) to immediate delivery or continuation of pregnancy²³. Only when such studies show a difference in outcome could one conclude that the CPR is of additional value to UA Doppler alone.

Conclusion

This review shows that calculating the CPR with MCA Doppler can improve on the accuracy of UA Doppler assessment in the prediction of adverse perinatal outcome in singleton pregnancies. This may be of particular importance in late FGR, although this could not be demonstrated with the available evidence. Clinical trials are needed to evaluate the effectiveness of the CPR for guiding clinical management in specific subgroups, such as late FGR fetuses.

ACKNOWLEDGMENTS

The authors thank Katja Jordanova, Marta Jozwiak and Rui Wang for helping with translations.

REFERENCES

- Royal College of Obstetricians and Gynaecologists. *The Investigation and Management of the Small-for-Gestational-Age Fetus*; Green-top Guideline No.31 (2nd edn). RCOG Press: London, 2013 (Minor revisions, January 2014).
- American College of Obstetricians and Gynecologists. ACOG Practice bulletin no. 134: fetal growth restriction. *Obstet Gynecol* 2013; **121**: 1122–1133.
- Lausman A, Kingdom J, Gagnon R, Basso M, Bos H, Crane J, Davies G, Delisle MF, Hudon L, Menticoglou S, Mundle W, Ouellet A, Pressey T, Pyllypuk C, Roggensack A, Sanderson F. Intrauterine growth restriction: screening, diagnosis, and management. *J Obstet Gynaecol Can* 2013; **35**: 741–757.

- Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. *Cochrane Database Syst Rev* 2013; **11**: CD007529.
- Morris RK, Say R, Robson SC, Kleijnen J, Khan KS. Systematic review and meta-analysis of middle cerebral artery Doppler to predict perinatal wellbeing. *Eur J Obstet Gynecol Reprod Biol* 2012; **165**: 141–155.
- Gramellini D, Folli MC, Raboni S, Vadora E, Meriardi A. Cerebral–umbilical Doppler ratio as a predictor of adverse perinatal outcome. *Obstet Gynecol* 1992; **79**: 416–420.
- Nassr AA, Abdelmagied AM, Shazly SA. Fetal cerebro-placental ratio and adverse perinatal outcome: systematic review and meta-analysis of the association and diagnostic performance. *J Perinat Med* 2016; **44**: 249–256.
- DeVore GR. The importance of the CPR in the evaluation of fetal well-being in SGA and AGA fetuses. *Am J Obstet Gynecol* 2015; **213**: 5–15.
- Dunn L, Sherrell H, Kumar S. Review: Systematic review of the utility of the fetal cerebroplacental ratio measured at term for the prediction of adverse perinatal outcome. *Placenta* 2017; **54**: 68–75.
- Meher S, Hernandez-Andrade E, Basheer SN, Lees C. Impact of cerebral redistribution on neurodevelopmental outcome in small-for-gestational-age or growth-restricted babies: a systematic review. *Ultrasound Obstet Gynecol* 2015; **46**: 398–404.
- Gordijn SJ, Beune IM, Thilaganathan B, Papageorghiou A, Baschat AA, Baker PN, Silver RM, Wynia K, Ganzevoort W. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol* 2016; **48**: 333–339.
- Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529–536.
- Rutter CM, Gatsonis CA. Regression methods for meta-analysis of diagnostic test data. *Acad Radiol* 1995; **2** (Suppl 1): S48–S56; discussion S65–S67, S70–S71.
- Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001; **20**: 2865–2884.
- 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**: 982–990.
- Mari G, Deter RL. Middle cerebral artery flow velocity waveforms in normal and small-for-gestational-age fetuses. *Am J Obstet Gynecol* 1992; **166**: 1262–1270.
- Arduini D, Rizzo G. Normal values of Pulsatility Index from fetal vessels: a cross-sectional study on 1556 healthy fetuses. *J Perinat Med* 1990; **18**: 165–172.
- Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and Presenting Results. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* Version 1.0, Deeks JJ, Bossuyt PM, Gatsonis C (eds), Chapter 10. The Cochrane Collaboration, 2010. Available from: <http://srda.cochrane.org/>.
- Boers KE, Vijgen SM, Bijlenga D, van der Post JA, Bekedam DJ, Kwee A, van der Salm PC, van Pampus MG, Spaanderman ME, de Boer K, Duvekot JJ, Bremer HA, Hasaart TH, Delemarre FM, Bloemenkamp KW, van Meir CA, Willekes C, Wijnen EJ, Rijken M, le Cessie S, Roumen FJ, Thornton JG, van Lith JM, Mol BW, Scherjon SA; DIGITAT study group. Induction versus expectant monitoring for intrauterine growth restriction at term: randomised equivalence trial (DIGITAT). *BMJ* 2010; **341**: c7087.
- Kazemier BM, Voskamp BJ, Ravelli AC, Pakrjt E, Groot CJ, Mol BW. Optimal timing of delivery in small for gestational age fetuses near term: a national cohort study. *Am J Perinatol* 2015; **30**: 177–186.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ (Clin Res Ed)* 2015; **351**: h5527.
- van 't Hooft J, Duffy JM, Daly M, Williamson PR, Meher S, Thom E, Saade GR, Alfirevic Z, Mol BW, Khan KS; Global Obstetrics Network (GONet). A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth. *Obstet Gynecol* 2016; **127**: 49–58.
- Bossuyt PM, Lijmer JG, Mol BW. Randomised comparisons of medical tests: sometimes invalid, not always efficient. *Lancet* 2000; **356**: 1844–1847.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:



Appendix S1 Search strategy implemented in electronic search of PubMed, EMBASE, the Cochrane library and ClinicalTrials.gov. Free text terms (Cochrane Library), MeSH (PubMed) and Emtree (EMBASE) controlled terms detailed.

Appendix S2 Reference details of 128 studies on prognostic accuracy of umbilical artery Doppler compared to cerebroplacental ratio and/or middle cerebral artery Doppler in prediction of adverse perinatal outcome included in systematic review and meta-analysis.

Appendix S3 QUADAS-2 quality evaluation of 128 included studies, showing risk of bias and concerns regarding applicability.

Appendix S4 Thresholds of middle cerebral artery Doppler (Table A) and of cerebroplacental ratio (Table B) used in included studies for the three different Doppler indices. References for studies that assessed normal ranges are provided.

Appendix S5 Reported sensitivities and specificities of middle cerebral artery Doppler (Table A) and cerebroplacental ratio (Table B) in prediction of adverse perinatal outcomes in studies investigating fetal growth restriction or small-for-gestational age at ≥ 32 weeks' gestation. Data are given as sensitivity (95% CI) or specificity (95% CI).

Table S1 Characteristics of 128 studies on prognostic accuracy of umbilical artery Doppler compared to cerebroplacental ratio and/or middle cerebral artery Doppler in prediction of adverse perinatal outcome included in systematic review and meta-analysis

Table S2 Heterogeneity analysis *P*-values of overall test performance of cerebroplacental ratio and middle cerebral artery Doppler. Fetal growth, term and timing of measurement were assessed for their impact on heterogeneity

Figure S1 Forest plots of sensitivity and specificity of middle cerebral artery Doppler (a) and cerebroplacental ratio (b) in prediction of different adverse perinatal outcomes in 128 included studies.



This article has been selected for Journal Club.

A slide presentation, prepared by Dr Maddalena Morlando, one of UOG's Editors for Trainees, is available online.

Chinese translation by Dr Yang Fang.
Spanish translation by Dr Rubén Darío Fernández.



Precisión en el pronóstico de resultados perinatales adversos mediante Doppler de la relación cerebroplacentaria y de la arteria cerebral media: revisión sistemática y metaanálisis

RESUMEN

Objetivo La evaluación ecográfica Doppler de la relación cerebroplacentaria (RCP) y la arteria cerebral media (ACM) se usa ampliamente como complemento del Doppler de la arteria umbilical (AU) para identificar fetos con riesgo de un resultado perinatal adverso. Sin embargo, los informes de las estimaciones de su precisión varían considerablemente. El objetivo de este estudio fue la revisión sistemática de las precisiones en el pronóstico mediante Doppler RCP y ACM para predecir resultados perinatales adversos y compararlas con el Doppler AU, a fin de identificar si la evaluación con Doppler RCP y ACM tiene un valor agregado al del Doppler AU.

Métodos Se realizaron búsquedas en PubMed, EMBASE, Cochrane Library y ClinicalTrials.gov, desde su inicio hasta junio de 2016, respecto a estudios sobre la precisión en el pronóstico del Doppler AU en comparación con Doppler RCP y/o ACM en la predicción de resultados perinatales adversos en embarazos con feto único de cualquier perfil de riesgo. El riesgo de sesgos y preocupaciones sobre la aplicabilidad se evaluó usando la herramienta QUADAS-2 (Evaluación de Calidad de Estudios de Precisión de Diagnósticos-2). Se realizó un metaanálisis respecto a múltiples resultados perinatales adversos. Utilizando modelos de metaregresión jerárquica de resumen de las características operativas del receptor, se comparó indirectamente la precisión del pronóstico del Doppler RCP versus ACM, y de forma directa la del Doppler RCP y ACM versus el Doppler AU.

Resultados La búsqueda identificó 4693 artículos, de los cuales se incluyeron 128 estudios (con 47748 mujeres). Se detectó un riesgo de sesgo o información subóptima en 120 de los 128 estudios (94%) y se encontró heterogeneidad sustancial, lo que limitó los análisis de subgrupos respecto al crecimiento fetal y la edad gestacional. Se observó una gran variación en las sensibilidades y especificidades reportadas, y en los valores umbral utilizados. El Doppler RCP superó al Doppler AU en la predicción de resultados adversos compuestos (tal y como los definen los estudios incluidos) ($P < 0,001$) y del parto de emergencia por sufrimiento fetal ($P = 0,003$), pero fue comparable al Doppler AU para los otros resultados. Los valores del Doppler ACM fueron significativamente peores que los del Doppler AU en la predicción de la puntuación de Apgar baja ($P = 0,017$) y del parto de emergencia por sufrimiento fetal ($P = 0,034$). El Doppler RCP superó al Doppler ACM en la predicción de resultados adversos compuestos ($P < 0,001$) y del parto de emergencia por sufrimiento fetal ($P = 0,013$).

Conclusión El cálculo del Doppler RCP junto con el Doppler ACM puede agregar valor a la evaluación con Doppler UA para la predicción de resultados perinatales adversos en mujeres con embarazo con feto único. Sin embargo, no está claro a qué subgrupo de mujeres embarazadas aplica esto. La efectividad de la RCP para guiar la atención clínica debe ser evaluada mediante pruebas clínicas.

脑—胎盘比值和大脑中动脉多普勒对不良围产结局的预测准确性: 系统评价和 meta 分析

目的: 超声多普勒评估脑—胎盘比值 (cerebroplacental ratio, CPR) 和大脑中动脉 (middle cerebral artery, MCA) 被广泛用作脐动脉 (umbilical artery, UA) 多普勒鉴别不良围产结局高危胎儿的辅助手段。然而, 文献报道的其准确性估计值变化很大。本研究的目的是系统评价 CPR 和 MCA 多普勒在预测不良围产结局时的预测准确性, 并与 UA 多普勒进行比较, 以证实 CPR 和 MCA 多普勒评估是否对 UA 多普勒具有辅助价值。

方法: 检索 PubMed、EMBASE、the Cochrane Library 和 ClinicalTrials.gov, 检索时间从建库至 2016 年 6 月, 检索在具有任何危险因素的单胎妊娠孕妇中, 将 UA 多普勒与 CPR 和/或 MCA 多普勒对预测不良围产结局的预测准确性进行比较的研究。采用 QUADAS-2 (诊断准确性研究的质量评价 2) 工具评估偏倚风险以及适用性。对多个不良围产结局进行 meta 分析。采用分层受试者工作特征 meta 回归模型, 对 CPR 与 MCA 多普勒的预测准确性进行间接比较, 对 CPR、MCA 多普勒与 UA 多普勒进行直接比较。

结果: 检索到 4693 篇文献, 其中纳入 128 项研究 (47 748 例孕妇)。128 项研究中有 120 项 (94%) 研究存在偏倚风险或次优报告, 并发现较大异质性, 从而限制了对胎儿生长和孕周进行亚组分析。报道的敏感性和特异性以及采用的临界值变化很大。在预测综合不良结局 (纳入研究对其定义) ($P < 0.001$) 以及由于胎儿窘迫急诊分娩 ($P = 0.003$) 方面, CPR 优于 UA 多普勒, 但在其他结局方面与 UA 多普勒相似。在预测 Apgar 评分较低 ($P = 0.017$) 以及由于胎儿窘迫急诊分娩 ($P = 0.034$) 方面, MCA 多普勒明显不如 UA 多普勒。在预测综合不良结局 ($P < 0.001$) 以及由于胎儿窘迫急诊分娩 ($P = 0.013$) 方面, CPR 优于 MCA 多普勒。

结论: 通过 MCA 多普勒计算 CPR 能够为采用 UA 多普勒评估预测单胎妊娠孕妇中不良围产结局提供辅助价值。然而, 还不清楚可以将其用于哪个孕妇亚组。需要进行临床试验, 评估 CPR 指导临床治疗的有效性。