DOI: 10.1111/1471-0528.13517
www.bjog.org

Systematic review

Antenatal magnetic resonance imaging versus ultrasound for predicting neonatal macrosomia: a systematic review and meta-analysis

GL Malin,[a] GJ Bugg,[a,b] Y Takwoingi,[c] JG Thornton,[a] NW Jones[a]

[a] School of Medicine, the University of Nottingham, Nottingham, UK [b] Department of Obstetrics, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham, UK [c] School of Health and Population Sciences, University of Birmingham, Birmingham, UK
Correspondence: G. Malin, Academic Department, School of Medicine, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK. Email gemma.malin@nottingham.ac.uk

Accepted 29 April 2015. Published Online 29 July 2015.

Background Fetal macrosomia is associated with an increased risk of adverse maternal and neonatal outcomes.

Objectives To compare the accuracy of antenatal two-dimensional (2D) ultrasound, three-dimensional (3D) ultrasound, and magnetic resonance imaging (MRI) in predicting fetal macrosomia at birth.

Search strategy Medline (1966–2013), Embase, the Cochrane Library and Web of Knowledge.

Selection criteria Cohort or diagnostic accuracy studies of women with a singleton pregnancy, who had third-trimester imaging to predict macrosomia (>4000 g, >4500 g or >90th or >95th centile).

Data collection and analysis Two reviewers screened studies, performed data extraction and assessed methodological quality. The bivariate model was used to obtain summary sensitivities, specificities and likelihood ratios.

Main results Fifty-eight studies (34 367 pregnant women) were included. Most were poorly reported. Only one study assessed 3D ultrasound volumetry. For predicting birthweight >4000 g or >90th centile, the summary sensitivity for 2D ultrasound (Hadlock) estimated fetal weight (EFW) >90th centile or >4000 g (29 studies) was 0.56 (95% CI 0.49–0.61), 2D ultrasound abdominal circumference (AC) >35 cm (four studies) was 0.80 (95% confidence interval [95% CI] 0.69–0.87) and MRI EFW (three studies) was 0.93 (95% CI 0.76–0.98). The summary specificities were 0.92 (95% CI 0.90–0.94), 0.86 (95% CI 0.74–0.93) and 0.95 (95% CI 0.92–0.97), respectively.

Conclusion There is insufficient evidence to conclude that MRI EFW is more sensitive than 2D ultrasound AC (which is more sensitive than 2D EFW); although it was more specific. Further primary research is required before recommending MRI EFW for use in clinical practice.

Keywords Estimated fetal weight, macrosomia, magnetic resonance imaging, pregnancy, three-dimensional ultrasound, two-dimensional ultrasound.

Tweetable abstract Systematic review of antenatal imaging to predict macrosomia. MRI EFW is more sensitive than ultrasound EFW.

Linked article: This article has journal club questions by EYL Leung, p. 89 in this issue. To view these visit http://dx.doi.org/10.1111/1471-0528.13518.

Please cite this paper as: Malin GL, Bugg GJ, Takwoingi Y, Thornton JG, Jones NW. Antenatal magnetic resonance imaging versus ultrasound for predicting neonatal macrosomia: a systematic review and meta-analysis. BJOG 2016;123:77–88.

Introduction

Macrosomia occurs as a result of excessive intrauterine fetal growth. A number of thresholds of birthweight have been used to define macrosomia, including >4000 g, >4500 g, >90th or >95th centile on a population nomogram.[1,2] Macrosomia is associated with an increased risk of shoulder dystocia and birth trauma, with associated adverse maternal and neonatal outcomes. These include maternal postpartum haemorrhage, third- and fourth-degree tears, and fractures, Erb's palsy and hypoxic injury to the infant.[3] These not only impact on the health of the individuals involved but represent a significant cost to the NHS, both for long-term care and settlement of litigation cases.[4]

Fetal macrosomia is associated with maternal diabetes mellitus (gestational or pre-existing) and obesity, both of which are increasing in incidence.[5] The risk of shoulder dystocia is significantly higher in infants of mothers with diabetes.[6]

Antenatal prediction of macrosomia is often inaccurate.[7] A variety of ultrasound measurements have been used for this purpose. A systematic review published in 2005 assessed the accuracy of two dimensional (2D) ultrasound biometry for prediction of macrosomia, and found that ultrasound was an overall poor predictor of fetal macrosomia, regardless of whether estimated fetal weight (EFW), computed from measurements of fetal head circumference (HC), abdominal circumference (AC) and femur length (FL), or fetal AC alone, was used.[8] As such, using ultrasound to assess the general antenatal population who are felt to be large for dates on clinical assessment is not recommended.[9]

Since publication of this review, three-dimensional (3D) ultrasound and magnetic resonance imaging (MRI) have become increasingly used in the assessment of the fetus in utero, and studies have examined their use in the estimation of fetal weight.[10,11] Therefore, the aim of this systematic review was to evaluate the predictive accuracy of 3D ultrasound and MRI for macrosomia at birth; to update the evidence on the accuracy of 2D ultrasound biometry; and to compare the accuracy of the three modalities.

Method

A protocol-driven systematic review was performed in accordance with published guidelines.[12,13] The review was registered with PROSPERO (CRD42013006127). The protocol is available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013006127. The reporting of the review meets the criteria specified in the PRISMA guidance.[14]

Literature search

We performed electronic searches from database inception until December 2013. We searched MEDLINE (1966 to December 2013), EMBASE (1980 to December 2013), the Cochrane Library (issue 12, 2013), the British Nursing Index (EBSCO) and MEDION, for relevant published articles. In order to identify 'grey' literature, OpenGrey, and Web of Science were also searched for relevant citations. In MEDLINE the search consisted of a combination of MESH headings (e.g. Pregnancy, Prenatal Ultrasonography), keywords (e.g. biometry, volumetry) and word variants using the Boelean operator 'OR' for capturing citations of the relevant text. These were combined using 'AND' with a combination of MESH headings (e.g. Fetal Macrosomia), keywords (e.g. birthweight, large for gestational age) and word variants to capture relevant outcomes. No language restrictions were applied. The MEDLINE search strategy is given in the Supplementary material (Appendix S1); this was adapted for use in other databases. Hand searching of recent major journals was also performed. A comprehensive database collating all citations was constructed using Reference Manager 12.0.

Study selection and data extraction

Initially, the database was scrutinised by two reviewers (GJB and GLM, 50% in duplicate) and full articles of all citations that were likely to meet the predefined selection criteria were obtained. Articles in languages other than English were translated. Following examination of full text articles, final inclusion or exclusion decisions were made by two reviewers (NWJ and GJB or GLM), adhering to the following criteria:

1 Population: Women with a singleton pregnancy
2 Index test: 2D or 3D ultrasound scan or MRI performed in the third trimester to detect fetal macrosomia. Several formulae are used for calculating EFW based on a combination of sonographic fetal measurements. Studies were included irrespective of the formula and threshold used to define macrosomia.
3 Reference standard (outcome): Birthweight >4000 g, >4500 g, >90th or >95th centile.
4 Study design: Diagnostic accuracy or cohort studies that allowed construction of 2 × 2 tables of the number of true positives, false positives, false negatives and true negatives. Studies with ten or fewer women were excluded because such studies were likely to provide unreliable estimates of test performance. Case–control studies were excluded because of their tendency to exaggerate the magnitude of test accuracy.[15,16]

All manuscripts were carefully examined to identify overlapping populations. Where this was the case, the most recent and complete manuscripts were selected. We applied no language restriction in the selection of studies. The reference lists of selected studies and review articles were checked and additional relevant articles were obtained. Data extraction was performed using a data collection sheet and was done in duplicate (NWJ and GJB or GLM). Data were extracted on study characteristics (including threshold values used), quality assessment criteria and test accuracy estimates, and were entered into an Excel spreadsheet. If results for multiple thresholds were reported, we constructed a separate 2 × 2 table for each threshold. In studies where data were felt to be relevant but 2 × 2 tables could not be constructed, we attempted to contact the authors. Disagreements between reviewers in the selection of studies and data extraction were resolved by consulting a third reviewer.

Quality assessment

The included studies were assessed for methodological quality using the QUADAS-2 checklist.[17] The checklist con-

sists of four domains: patient selection, index test, reference standard and flow and timing. Based on a number of signalling questions, each of the four domains are assessed for risk of bias but only the first three domains are assessed for applicability concerns. Patient selection was considered to have a high risk of bias if anything other than a consecutively or randomly recruited study population was reported. If the threshold of the index test used to predict macrosomia was not prespecified, a high risk of bias was reported. We defined an appropriate interval between the index test (antenatal scan) and reference standard (birthweight) as <7 days. If the time lapse was greater, we assigned a high risk of bias in the 'flow and timing' domain. Where studies were inadequately reported, making it not possible to make a clear judgement about risk of bias or applicability concern, the category was assigned an 'unclear risk' or 'unclear concern'. We planned to perform sensitivity analyses where possible by excluding studies at 'high or unclear risk of bias' in a domain.

Data synthesis

To assess the diagnostic accuracy of a test, we used the 2 × 2 tables to calculate sensitivity, specificity, positive and negative likelihood ratios, and their 95% confidence intervals (CI). We plotted estimates of sensitivity and specificity in receiver operating characteristic space and on forest plots to explore between study differences in estimates of test performance.

For studies that used the same index test and reference standard at the same threshold, we performed meta-analysis using the bivariate model to jointly synthesise sensitivity and specificity.[18,19] The bivariate model includes random effects that allow for between-study variation in sensitivity and specificity, as well as a correlation parameter that allows for the trade-off in sensitivity and specificity across studies. We used parameter estimates from the model to derive likelihood ratios with their 95% confidence intervals.[20] Given the complexity of the bivariate model, where few studies were available, we simplified the model by removing the correlation parameter or assuming fixed effects for sensitivity and/or specificity. We included studies that defined macrosomia using either birthweight >90th centile or >4000 g in the same meta-analysis because both are generally considered to be similar. However, we also performed subgroup analyses considering each definition independently.

To compare the accuracy of the tests, we added a covariate for test type to the bivariate model to assess its effect on sensitivity and specificity. The statistical significance of the difference in test performance was assessed using a likelihood ratio test comparing models with and without the covariate terms. We investigated the type of population (high risk versus low risk) as a potential source of heterogeneity by adding a covariate to the bivariate model. We classified a study population as high risk if the women in question had a higher risk of fetal macrosomia than the general obstetric population, including pre-existing or gestational diabetes, obesity or postmaturity (>40 weeks of gestation). Low risk was classified as women without these conditions, or an unselected obstetric population.

Determinants of publication bias are not well understood for test accuracy studies. Commonly used methods for assessing publication bias are not appropriate for test accuracy reviews.[18] We planned to use the recommended approach of Deeks et al.[21] if there was minimal heterogeneity because like all other approaches, this approach has low power for detecting funnel plot asymmetry when there is heterogeneity.[21] All analyses were performed in Stata version 13.0 (StataCorp, College Station, TX, USA) and meta-analyses were performed using the Stata xtmelogit command. We used Review Manager (version 5.3; Copenhagen; The Nordic Cochrane Center, The Cochrane Collaboration, 2014) to generate forest plots and summary receiver operating characteristic plots.

Results

We identified 4140 citations, of which 459 were considered for full text review. Of these, 401 articles were excluded for reasons shown in Figure 1. Fifty-eight studies involving 34 367 pregnant women were included.[10,11,22–77] The characteristics of included studies are given in the Supplementary material (Table S1). The majority of studies (97%) reported 2D ultrasound parameters, with only one study providing data on the accuracy of 3D ultrasound volumetry. There were five studies with data on antenatal MRI volumetry. One of these was performed in our centre.[51] Of the remaining four studies, two used the same study population.[11,78] We therefore included only the most complete data set.[11] Of the remaining two studies, one contained fewer than five women and was therefore excluded.[79] There were concerns that the population in the second study did not meet our inclusion criteria as women undergoing termination of pregnancy at >20 weeks of gestation were also included, and scans were not restricted to the third trimester as we prespecified.[46] We attempted to contact the authors to obtain the relevant data, but received no response. As there was paucity of data on MRI, we included this study, which compared MRI and 2D ultrasound in the same women with the intention of performing a sensitivity analysis excluding the study.

Quality assessment

Methodological quality varied between studies but risk of bias was generally unclear because of poor reporting. The

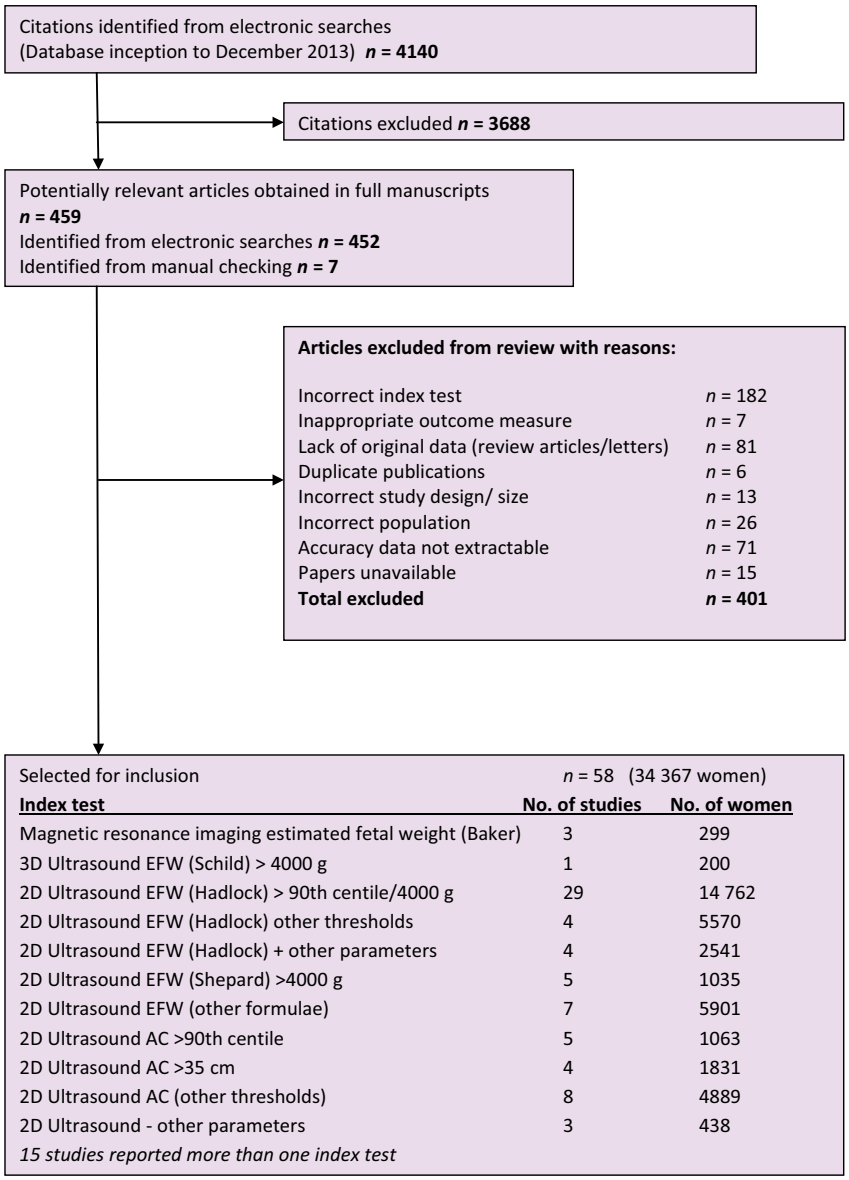

Figure 1. Study selection process for systematic review of the diagnostic accuracy of antenatal ultrasound and MRI scan for fetal macrosomia at birth.

summary of the quality assessment of the included studies is shown in Figure 2. None of the studies was scored as low risk of bias in all domains. Two studies[22,63] were scored as either high or unclear risk of bias in all four domains. Most of the studies were scored as unclear risk of bias in the patient selection domain (76%) and reference standard domain (95%). In contrast, most of the studies were scored as low risk of bias in the index test domain (79%). In the flow and timing domain, 45% of studies were scored as high risk of bias. As most studies were scored high or unclear risk of bias in all domains except the index test domain, we did not perform sensitivity analyses. Most studies (79%) had low concern regarding applicability across all three domains. One study[46] that gave particular concern regarding applicability has been described above.

Accuracy of 2D ultrasound parameters for predicting fetal macrosomia

Thirty-one studies estimated fetal weight using any Hadlock formula at a threshold of EFW >90th centile or >4000 g to predict birthweight defined at the same thresholds. Two study populations[24,56] were duplicated and were therefore excluded from the meta-analysis of studies using

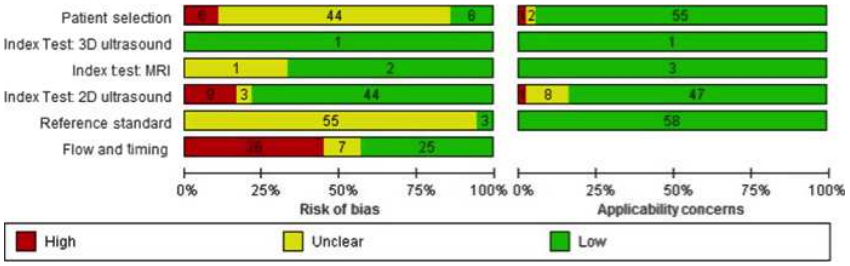

Figure 2. Summary of methodological quality assessment of risk of bias and applicability concerns presented for each domain as percentages across all included studies. The numbers on the bar for each domain represents the number of studies that were scored as high, unclear or low risk of bias or applicability concern.

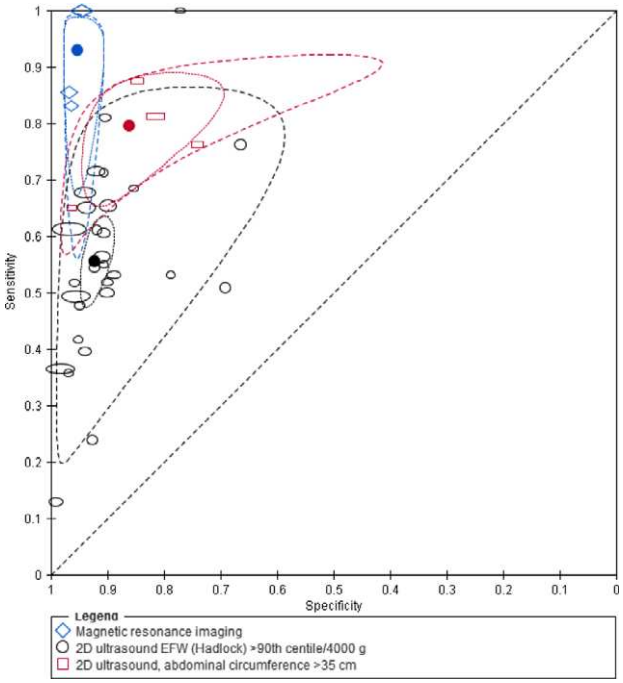

Figure 3. Summary receiver operating characteristic plot of MRI, 2D ultrasound EFW using any Hadlock formula at threshold EFW >90th centile or >4000 g, and AC >35 cm for prediction of macrosomia. The symbol for each test represents the pair of sensitivity and specificity from a study. The symbols are scaled according to sample size. The solid circles represent the summary sensitivity and specificity for each test. The summary points are surrounded by 95% confidence regions (dotted line) and 95% prediction regions (dashed line).

either EFW >90th centile or >4000 g, but were included in the subgroup analyses that considered EFW >90th centile and EFW >4000 g separately. The sensitivities from the 29 studies (2085 cases, 14 762 women) ranged between 0.13 and 1.00, and the specificities ranged between 0.66 and 0.99 (see Supplementary material, Figure S1). The median pretest probability (prevalence) of fetal macrosomia, calculated using the 29 studies, was 11% (interquartile range 19–26%). Substantial heterogeneity was observed as shown by the extent of the 95% prediction region around the summary point on the summary receiver operating characteristic plot (Figure 3), The summary sensitivity and specificity of 2D ultrasound EFW were 0.56 (95% CI 0.49–0.61) and 0.92 (95% CI 0.90–0.94), respectively (Table 1). The summary positive and negative likelihood ratios were 7.2 (5.5–9.4) and 0.48 (0.42–0.55).

Results obtained from the subgroup analyses of EFW >90th centile to predict birthweight >90th centile, and EFW >4000 g to predict birthweight >4000 g were similar to those from the combined analysis (see Supplementary material, Table S2). A sensitivity analysis excluding the study by Kacem et al.[46] also did not change the result. A sensitivity analysis including only the five studies[43,49,59,62,73] that used any Hadlock formula incorporating HC, AC and FL to compute estimated fetal weight gave similar results to the analysis that included studies using any version of the Hadlock formula. Investigation of type of population as a potential source of heterogeneity showed no evidence of a difference in sensitivity ($P = 0.9$) and specificity ($P = 0.4$) between low-risk and high-risk populations (see Supplementary material, Table S3). We did not assess publication bias because of the observed heterogeneity.

Estimates of sensitivity and specificity from studies that used the Hadlock formula at other thresholds, the Hadlock formula plus other parameters, or other EFW formulae are given in Figure S2 (see Supplementary material). For AC >35 cm (four studies), the summary sensitivity and specificity were 0.80 (95% CI 0.69–0.87) and 0.86 (95% CI 0.74–0.93), respectively. Figure S3 (see Supplementary material) shows the estimates of sensitivity and specificity for studies that used other AC thresholds. Figure S4 (see Supplementary material) shows the estimates for studies that used other 2D ultrasound parameters. Where meta-analyses were possible, the summary estimates obtained are presented in Table 1. The summary sensitivity and specificity of 2D ultrasound using the Shepard formula were similar to the summary estimates for the Hadlock formula.

Table 1. Accuracy of 2D ultrasound, 3D ultrasound and MRI for predicting fetal macrosomia

Index test and threshold	Reference standard	Number of studies (Cases/total number of women)	Sensitivity (95% CI)	Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)
Comparison of 2D ultrasound EFW (any Hadlock formula >90th centile or >4000 g) and MRI						
2D ultrasound EFW (any Hadlock) >90th centile or >4000 g	Birthweight >90th centile or >4000 g	29 (2085/14762)	0.56 (0.49–0.62)	0.92 (0.90–0.94)	7.2 (5.5–9.4)	0.48 (0.42–0.55)
2D ultrasound AC >35 cm	Birthweight >90th centile or >4000 g	4 (113/1831)	0.80 (0.69–0.87)	0.86 (0.74–0.93)	5.8 (3.1–10.6)	0.24 (0.16–0.35)
MRI EFW >90th centile or >4000 g	Birthweight >90th centile or >4000 g	3 (41/299)	0.93 (0.76–0.98)	0.95 (0.92–0.97)	20.0 (9.6–41.7)	0.07 (0.02–0.28)
Test for difference in sensitivity or specificity of 2D ultrasound EFW, 2D ultrasound AC >35 cm and MRI: $P = 0.0002$*						
3D ultrasound and other 2D ultrasound estimated fetal weight formulae or parameters						
3D ultrasound EFW (Schild) >4000 g	Birthweight >4000 g	1 (33/200)	0.42 (0.26–0.61)	0.98 (0.95–1.00)	23.6 (7.2–77.6)	0.59 (0.44–0.79)
2D ultrasound EFW (Ott) >4000 g	Birthweight >4000 g	2 (61/239)	0.44 (0.32–0.57)	0.96 (0.91–0.98)	9.8 (5.0–19.5)	0.58 (0.47–0.73)
2D ultrasound EFW (Woo) >4000 g	Birthweight >4000 g	2 (686/3067)	0.54 (0.50–0.58)	0.90 (0.89–0.92)	5.6 (5.0–6.4)	0.51 (0.47–0.55)
2D ultrasound EFW (Shepard) >4000 g	Birthweight >90th centile or >4000 g	5 (218/1035)	0.55 (0.41–0.68)	0.89 (0.84–0.93)	5.1 (3.9–6.7)	0.50 (0.39–0.66)
2D ultrasound AC >36 cm	Birthweight >90th centile or >4000 g	2 (60/208)	0.67 (0.54–0.77)	0.90 (0.84–0.94)	6.6 (4.0–10.8)	0.37 (0.26–0.53)
2D ultrasound AC >90th centile	Birthweight >90th centile or >4000 g	5 (167/1063)	0.75 (0.61–0.85)	0.89 (0.74–0.96)	7.4 (3.1–17.7)	0.28 (0.19–0.41)
2D ultrasound AC >37 cm**	Birthweight >4000 g	2 (46/192)	–	–	–	–

**P*-value from the likelihood ratio test that assessed the statistical significance of the difference in sensitivity or specificity between the three tests by comparing the model without covariate terms with the model that included covariate terms for type of test.

***Meta-analysis was not performed because a fixed effect model was used when there were only two studies. A model that does not account for between-study variability would not be appropriate given the observed heterogeneity.

Accuracy of 3D ultrasound for predicting fetal macrosomia

Only one study[10] reported the accuracy of 3D ultrasound volumetry EFW to predict birthweight >4000 g (see Supplementary material, Figure S1). The sensitivity was 0.42 (95% CI 0.26–0.61) and specificity was 0.98 (95% CI 0.95–1.00).

Accuracy of MRI for predicting fetal macrosomia

Three studies (41 cases, 299 women) using fetal volumes and the formula reported by Baker et al.[80] to estimate fetal weight were included in the meta-analysis (Figure 3, and see Supplementary material, Figure S1). The median pretest probability (prevalence) of fetal macrosomia was 11% (range 17–18%). Two studies used a threshold of >4000 g, and one used >90th centile. The summary sensitivity of MRI was 0.93 (95% CI 0.76–0.98) and summary specificity was 0.95 (95% CI 0.92–0.97). The positive and negative likelihood ratios were 20.0 (95% CI 9.6–41.7) and 0.07 (95% CI 0.02–0.28). When we excluded the study that did not match our inclusion criteria exactly, the sensitivity was 0.85 (95% CI 0.62–0.95) and specificity was 0.97 (95% CI 0.90–0.99).

Comparison of the accuracy of 2D ultrasound and MRI

Based on 34 studies, there was a significant difference ($P = 0.0002$) in the performance of MRI (three studies), 2D ultrasound (Hadlock) EFW (29 studies), and AC >35 cm (four studies) for predicting macrosomia. The sensitivity of MRI was significantly superior to that of 2D ultrasound EFW ($P = 0.001$), and despite a 13% difference in sensitivity between MRI and AC >35 cm, the difference was not statistically significant ($P = 0.12$). The specificity of MRI was significantly higher than that of 2D ultrasound AC >35 cm ($P = 0.02$) but there was no evidence of a difference in specificity between MRI and 2D ultrasound EFW ($P = 0.11$). 2D ultrasound AC >35 cm was more sensitive than 2D ultrasound EFW ($P = 0.003$), but less specific ($P = 0.012$).

There were no studies that compared 2D ultrasound AC and MRI in the same population. The findings of the comparison of meta-analyses between 2D ultrasound EFW and MRI EFW were consistent with the findings of the two studies that compared these tests in the same population. Zaretsky et al.[11] found that 2D ultrasound (Hadlock) EFW >4000 g had a sensitivity of 0.35 (95% CI 0.13–0.65) and specificity of 0.97 (95% CI 0.89–1.00) compared with MRI EFW >4000 g with a sensitivity of 0.86 (95% CI 0.57–0.98) and specificity of 0.97 (95% CI 0.89–1.00).

Kacem et al.[46] found that ultrasound EFW (Hadlock) >4000 g had sensitivity of 0.81 (95% CI 0.58–0.95) and specificity of 0.90 (95% CI 0.85–0.94), while MRI EFW >4000 g had sensitivity of 1.00 (95% CI 0.84–1.00) and specificity of 0.95 (95% CI 0.90–0.98).

For illustration purposes, given a pretest probability of 17% (median from the 34 studies), the findings imply that in a hypothetical cohort of 1000 pregnant women, 170 babies will be born with birthweight >90th centile or >4000 g. Of the 170 macrosomic babies, 2D ultrasound EFW will miss 75, AC >35 cm will miss 34 while MRI will only miss 12 babies. Of the 830 babies without macrosomia, 2D ultrasound EFW will incorrectly identify 66 and AC will incorrectly identify 116 as macrosomic whereas MRI EFW will incorrectly identify 42 babies as macrosomic. Using the likelihood ratios and the same pretest probability, the post-test probability of having macrosomia is 9.0 and 4.7% for a negative test result for 2D ultrasound EFW and AC >35 cm respectively, and 60 and 54% for a positive test result. However, the post-test probability of having macrosomia is 1.4% if MRI EFW is negative and 80% if MRI EFW is positive for macrosomia.

Discussion

Main findings

The majority of studies reported 2D ultrasound EFW, calculated with any Hadlock formula, for a birthweight threshold of >4000 g or >90th centile. Our results show that this test has reasonable specificity (i.e. low false-positive rate) but poor sensitivity (high false-negative rate). The likelihood ratios indicate that a positive result may rule in macrosomia, but a negative result does not rule it out. The Perinatal Institute recommend using a Hadlock formula including HC, AC and FL for estimating fetal weight.[81] When we performed sensitivity analysis limited to studies using the same parameters and citing the same reference[82] the results were similar.

MRI volumetry to estimate fetal weight appeared to be much more sensitive than 2D ultrasound EFW for predicting fetal macrosomia. However, these results were based on few studies and small numbers.

Strengths and limitations

The strength of our review and the validity of our inferences lie in the methodology and the volume of evidence. We have complied with existing guidelines for conducting[13] and reporting systematic reviews.[14] We have used recommended techniques for performing and interpreting meta-analysis.[19,83] An extensive literature search was performed without language restrictions.

There are several limitations to our review. Despite extensive searches, there was a paucity of data on 3D ultrasound. Only one study was identified where data could be extracted to populate a 2 × 2 table. Six other studies were

identified through our searches.[84–89] However, none of them looked specifically at using 3D ultrasound to predict macrosomia, and often included small numbers of macrosomic infants. When a comparison of 2D and 3D ultrasound was made for general fetal weight estimation, one study found that combining fetal thigh volume with 2D biometry gave a higher proportion of estimated weights within 5% of actual birthweight than 2D ultrasound using the Hadlock formula alone (70 versus 40%).[89] However, three other studies found no difference in the accuracy of weight estimation between 2D and 3D ultrasound.[84–89]

Only three MRI studies were included. One of these did not strictly meet our inclusion criteria.[46] We explored the impact of including this study by performing a sensitivity analysis. Excluding the study did not affect the summary estimates for 2D ultrasound EFW, but the summary sensitivity of MRI was reduced. However, the sensitivity of MRI for predicting fetal macrosomia was still greater than that of ultrasound, and we are confident that our conclusions are valid.

Differences may exist between the study populations that were combined in the meta-analyses, and the techniques used for the index tests and reference standard may have differed. There are several different formulae reported by Hadlock in calculating estimated fetal weight.[82,90–93] It was not always possible to determine from the primary reports exactly which formula was used. Although all studies cited a reference, multiple formulae were reported in Hadlock's original reports. This may result in some variability within the meta-analysis groups including 'any' Hadlock formula, because variations in accuracy between these formulae have been reported.[59,61] In order to address this, we performed a sensitivity analysis including only studies that cited the formula considered to be the most accurate. The results were comparable with the overall analysis.

Another limitation of this review is that by only analysing sensitivity and specificity according to prespecified thresholds of birthweight, we did not assess the mean absolute error of the individual techniques for birthweight as a continuous variable. We were limited by the data reported in the primary studies, which tended to report outcomes according to prespecified thresholds. Individual patient data meta-analysis could be performed to address this, but was considered beyond the scope of the current project.

A comparison of the accuracy of different tests using studies that have not compared the tests in the same study population may be prone to confounding.[94] However, only two studies compared MRI and 2D ultrasound EFW in the same population. The individual study results were consistent with the findings of the meta-analysis that used all available studies to compare both tests.

Comparison with other studies

We are not aware of any other systematic reviews examining the accuracy of antenatal MRI or 3D ultrasound for macrosomia at birth. For 2D ultrasound EFW, our results were similar to those published by Coomarasamy et al. in 2005.[8] They concluded that there was a lack of precision in 2D ultrasound in the prediction of fetal macrosomia, and that a positive test was more accurate for ruling in macrosomia, than a negative test was for ruling it out.[8] We identified and included a number of studies (with larger numbers of women) published subsequent to this review. We found that AC >35 cm appeared to be more sensitive that ultrasound EFW >90th or 4000 g, but less specific, and therefore overall a poor predictive test. Our results also agree with the findings of the NICE Antenatal Care guideline.[9]

Implications for clinical practice

MRI appears to show promise as an accurate antenatal test for predicting fetal macrosomia at birth. This has the potential to improve care for both women at high risk of macrosomia (including those with diabetes and obesity) and low-risk women. MRI is becoming increasingly available and is known to be safe in pregnancy, making its use for estimating fetal weight a realistic prospect. However, there are several considerations before MRI can be used for this purpose. This meta-analysis included very few studies with small numbers of women receiving MRI compared with evaluations of ultrasound. Therefore, there is still uncertainty regarding the accuracy of MRI, and further research is required before it can be considered for clinical practice. The cost and feasibility of performing MRI on a large scale must also be considered. There is a potential role for using AC as a screening test to determine which women should undergo fetal MRI EFW.

Recommendations

Further research is required, comparing MRI with 2D and 3D ultrasound in the same population. Given the findings in our study that 2D ultrasound AC >35 cm appeared to have a higher sensitivity (i.e. low false negative rate) than 2D ultrasound EFW a diagnostic strategy worthy of exploration is a two-stage approach using AC >35 cm to screen women. Those that screen positive would go on to have an MRI. If the accuracy of this approach is proven to be comparable to that of MRI alone, this approach is likely to be cheaper and more feasible, therefore applicable to a wider antenatal population.

The other uncertainty regarding clinical practice in this field is not only how best to predict fetal weight antenatally, but also how that information should be used. While macrosomia, defined according to the thresholds used in

this review, has been associated with adverse outcomes, a large number of macrosomic babies deliver without problems, and shoulder dystocia and trauma occur where babies are average weight.[95] Although this review did not consider labour outcomes, these must be considered in the evaluation of diagnosis–treatment pathways involving MRI estimated fetal weight, and whether early induction of labour or caesarean section is cost-effective in reducing adverse outcomes.

Conclusion

Using antenatal MRI to estimate fetal weight shows promising accuracy for the prediction of macrosomia at birth. However, further research is required before this technique can be applied in clinical practice.

Disclosure of interest

None declared. Completed disclosure of interests form available to view online as supporting information.

Contribution to authorship

GLM conceived the review, carried out data extraction, analysis and interpretation of data and drafted the article, and is responsible for the integrity of the work as a whole. GJB carried out data extraction and interpretation of data, revised the article critically for intellectual content and approved the final draft for publication. YT carried out statistical analysis and interpretation of the data, revised the article critically for intellectual content and approved the final draft for publication. JT assisted with interpretation of the data, revised the article critically for intellectual content and approved the final draft for publication.

N W Jones carried out data extraction and interpretation, revised the article critically for intellectual content and approved the final draft for publication.

Details of ethics approval

Ethical approval was not sought, as this was evidence synthesis of previously published work.

Funding

There was no external funding for this work.

Acknowledgements

We thank Dr Karla Hemming for her helpful suggestions for improving the manuscript.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Forest plot of MRI, 3D ultrasound and 2D ultrasound estimated fetal weight (using Hadlock formula) and abdominal circumference for prediction of macrosomia.

Figure S2. Forest plot of 2D ultrasound estimated fetal weight (EFW) using other Hadlock thresholds, Hadlock formula plus other parameters or other EFW formulae to predict macrosomia.

Figure S3. Forest plot of 2D ultrasound using abdominal circumference (all thresholds except >35 cm) to predict macrosomia.

Figure S4. Forest plot of other 2D ultrasound parameters to predict macrosomia.

Table S1. Characteristics of included studies in systematic review of the diagnostic accuracy of antenatal ultrasound or MRI scan for macrosomia at birth

Table S2. Subgroup and sensitivity analyses of 2D ultrasound EFW (any Hadlock formula) > 90th centile or > 4000 g

Table S3. Investigation of heterogeneity—effect of type of population on diagnostic accuracy of 2D ultrasound EFW (any Hadlock formula) >90th centile or >4000 g

Appendix S1. Medline search strategy for systematic review of the accuracy of antenatal ultrasound and magnetic resonance imaging for macrosomia at birth

Data S1. Powerpoint slides summarising the study. ■

References

1 Allen K, Wallace SVF. Fetal macrosomia. *Obstet Gynecol Reprod Biol* 2013;6:185–8.
2 Lowe LP, Metzger BE, Dyer AR. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study. Associations of maternal A1C and glucose with pregnancy outcomes. *Diabetes Care* 2012;35:574–80.
3 King JR, Korst LM, Miller DA, Ouzounian JG. Increased composite maternal and neonatal morbidity associated with ultrasonographically suspected fetal macrosomia. *J Matern Fetal Neonatal Med* 2012; 25:1953–9.
4 NHS Litigation Authority. Learning from of maternity claims. 2012. London, NHSLA. http://www.nhsla.com/CurrentActivity/Documents/Learning%20from%20Maternity%20Claims.pdf. Accessed 20 October 2014.
5 Young BC, Ecker JL. Fetal macrosomia and shoulder dystocia in women with gestational diabetes: risks amenable to treatment? *Curr Diab Rep* 2013;13:12–8.
6 Nesbitt TS, Gilbert WM, Herrchen B. Shoulder dystocia and associated risk factors with macrosomic infants born in California. *Am J Obstet Gynecol* 1998;179:476–80.
7 Royal College of Obstetricians and Gynaecologists. *Shoulder Dystocia (Green Top guideline No.42)*. London: RCOG; 2012. [www.rcog.org.uk/globalassets/documents/guidelines/gtg42_25112013.pdf] Accessed 7 March 2015.
8 Coomarasamy A, Connock M, Thornton J, Khan KS. Accuracy of ultrasound biometry in the prediction of macrosomia: a systematic quantitative review. *BJOG* 2005;112:1461–6.

9 National Institute for Health and Care Excellence (NICE). Antenatal care. 2011. [www.nice.org.uk/guidance/CG62/chapter/introduction]. Accessed 30 October 2014.

10 Hasenoehrl G, Pohlhammer A, Gruber R, Staudach A, Steiner H. Fetal weight estimation by 2D and 3D ultrasound: comparison of six formulas. *Ultraschall Med* 2009;30:585–90.

11 Zaretsky MV, Reichel TF, McIntire DD, Twickler DM. Comparison of magnetic resonance imaging to ultrasound in the estimation of birth weight at term. *Am J Obstet Gynecol* 2003;189:1017–20.

12 Akers J, Aguiar-Ibáñez R, Baba-Akbari Sari A, Beynon S, Booth A. Systematic Reviews: CRD's guidance for undertaking reviews in health care. York UK, NHS Centre for Reviews and Dissemination, University of York; 2009.

13 Deeks J, Bossuyt PM. Guide to the contents of a Cochrane Diagnostic Test Accuracy protocol. Cochrane Handbook DTA reviews. 2013 [http://srdta.cochrane.org/handbook-dta-reviews]. Accessed 2 March 2015.

14 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097. doi:10.1371/journal.pmed1000097.

15 Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JHP, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061–6.

16 Burch J, St John J, Duffy S, Smith S, Soares-Weiser K, Kleijnen J, et al. Should data from diagnostic case–control studies be included in systematic reviews alongside diagnostic cohort studies?. 2014. University of York, Centre for Reviews and Dissemination. [www.york.ac.uk/inst/crd/Posters/Should%20data%20from%20diagnostic%20casecontrol%20studies%20be%20included%20in%20systematic%20reviews%20alongside%20diagnostic%20cohort%20studies.pdf]. Accessed 17 November 2014.

17 Whiting P, Rutjes AW, Westwood ME, Mallett S, Deeks J, Reitsma JB. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Int Med* 2011;155:529–36.

18 Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: analysing and presenting results. In: Deeks JJ, Gatsonis C, Harbord RM, editors. *CochraneHandbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. The Cochrane Collaboration; 2010 [http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/Chapter%2010%20-%20Version%201.0.pdf]. Accessed 2 March 2015.

19 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;58:982–90.

20 Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. *Stat Med* 2008;27:687–97.

21 Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882–93.

22 Al-Inany H, Alaa N, Momtaz M, Abdel BM. Intrapartum prediction of macrosomia: accuracy of abdominal circumference estimation. *Gynecol Obstet Invest* 2001;51:116–9.

23 Ben-Haroush A, Melamed N, Mashiach R, Meizner I, Yogev Y. Use of the amniotic fluid index combined with estimated fetal weight within 10 days of delivery for prediction of macrosomia at birth. *J Ultrasound Med* 2008;27:1029–32.

24 Ben-Haroush A, Yogev Y, Hod M, Bar J. Predictive value of a single early fetal weight estimate in normal pregnancies. *Eur J Obstet Gynecol Reprod Biol* 2007;130:187–92.

25 Ben-Haroush A, Yogev Y, Mashiach R, Hod M, Meisner I. Accuracy of sonographic estimation of fetal weight before induction of labor in diabetic pregnancies and pregnancies with suspected fetal macrosomia. *J Perinat Med* 2003;31:225–30.

26 Benacerraf BR, Gelman R, Frigoletto FD Jr. Sonographically estimated fetal weights: accuracy and limitation. *Am J Obstet Gynecol* 1988;159:1118–21.

27 Benson CB, Doubilet PM, Saltzman DH. Sonographic determination of fetal weights in diabetic pregnancies. *Am J Obstet Gynecol* 1987;156:441–4.

28 Bethune M, Bell R. Evaluation of the measurement of the fetal fat layer, interventricular septum and abdominal circumference percentile in the prediction of macrosomia in pregnancies affected by gestational diabetes. *Ultrasound Obstet Gynecol* 2003;22:586–90.

29 Bian XM. Choice of ultrasound fetal weight estimation formulae. *Zhonghua Yi Xue Za Zhi* 1992;72:677–9.

30 Chauhan SP, Sullivan CA, Magann EF, Perry KG Jr, Roberts WE, Morrison JC. Estimate of birth weight among post-term pregnancy: clinical versus sonographic. *J Matern Fetal Med* 1994;3:208–11.

31 Chauhan SP, Hendrix NW, Magann EF, Morrison JC, Kenney SP, Devoe LD. Limitations of clinical and sonographic estimates of birth weight: experience with 1034 parturients. *Obstet Gynecol* 1998;91:72–7.

32 Chauhan SP, Parker D, Shields D, Sanderson M, Cole JH, Scardo JA. Sonographic estimate of birth weight among high-risk patients: feasibility and factors influencing accuracy. *Am J Obstet Gynecol* 2006;195:601–6.

33 Chen CP, Chang FM, Chang CH, Lin YS, Chou CY, Ko HC. Prediction of fetal macrosomia by single ultrasonic fetal biometry. *J Formos Med Assoc* 1993;92:248.

34 Chervenak JL, Divon MY, Hirsch J, Girz BA, Langer O. Macrosomia in the postdate pregnancy: is routine ultrasonographic screening indicated? *Am J Obstet Gynecol* 1989;161:753–6.

35 Cohen JM, Hutcheon JA, Kramer MS, Joseph KS, Abenhaim H, Platt RW. Influence of ultrasound-to-delivery interval and maternal-fetal characteristics on validity of estimated fetal weight. *Ultrasound Obstet Gynecol* 2010;35:434–41.

36 Colman A, Maharaj D, Hutton J, Tuohy J. Reliability of ultrasound estimation of fetal weight in term singleton pregnancies. *N Z Med J* 2006;119:U2146.

37 Combs CA, Rosenn B, Miodovnik M, Siddiqi TA. Sonographic EFW and macrosomia: is there an optimum formula to predict diabetic fetal macrosomia? *J Matern Fetal Med* 2000;9:55–61.

38 De Reu PA, Smits LJ, Oosterbaan HP, Nijhuis JG. Value of a single early third trimester fetal biometry for the prediction of birth weight deviations in a low risk population. *J Perinat Med* 2008;36:324–9.

39 Freire DMC, Cecatti JG, Paiva CSM. Correlation between estimated fetal weight by ultrasound and neonatal weight [Portuguese]. *Rev Bras Ginecol Obstet* 2010;32:4–10.

40 Gilby JR, Williams MC, Spellacy WN. Fetal abdominal circumference measurements of 35 and 38 cm as predictors of macrosomia. A risk factor for shoulder dystocia. *J Reprod Med* 2000;45:936–8.

41 Hendrix NW, Grady CS, Chauhan SP. Clinical vs. sonographic estimate of birth weight in term parturients. A randomized clinical trial. *J Reprod Med* 2000;45:317–22.

42 Henrichs C, Magann EF, Brantley KL, Crews JH, Sanderson M, Chauhan SP. Detecting fetal macrosomia with abdominal circumference alone. *J Reprod Med* 2003;48:339–42.

43 Holcomb WL Jr, Mostello DJ, Gray DL. Abdominal circumference vs. estimated weight to predict large for gestational age birth weight in diabetic pregnancy. *Clin Imaging* 2000;24:1–7.

44 Humphries J, Reynolds D, Bell-Scarbrough L, Lynn N, Scardo JA, Chauhan SP. Sonographic estimate of birth weight: relative accuracy of sonographers versus maternal-fetal medicine specialists. *J Matern Fetal Neonatal Med* 2002;11:108–12.

45 Johnstone FD, Prescott RJ, Steel JM, Mao JH, Chambers S, Muir N. Clinical and ultrasound prediction of macrosomia in diabetic pregnancy. *BJOG* 1996;103:747–54.

46 Kacem Y, Cannie MM, Kadji C, Dobrescu O, Lo Zito L, Ziane S. Fetal weight estimation: comparison of two-dimensional US and MR imaging assessments. *Radiology* 2013;267:902–10.

47 Kayem G, Grange G, Breart G, Goffinet F. Comparison of fundal height measurement and sonographically measured fetal abdominal circumference in the prediction of high and low birth weight at term. *Ultrasound Obstet Gynecol* 2009;34:566–71.

48 Landon MB, Mintz MC, Gabbe SG. Sonographic evaluation of fetal abdominal growth: predictor of the large-for-gestational-age infant in pregnancies complicated by diabetes mellitus. *Am J Obstet Gynecol* 1989;160:115–21.

49 Levine AB, Lockwood CJ, Brown B, Lapinski R, Berkowitz RL. Sonographic diagnosis of the large for gestational age fetus at term: does it make a difference? *Obstet Gynecol* 1992;79:55–8.

50 Loetworawanit R, Chittacharoen A, Sututvoravut S. Intrapartum fetal abdominal circumference by ultrasonography for predicting fetal macrosomia. *J Med Assoc Thai* 2006;4:S60–4.

51 Malin G, Angbalan D, Bugg G, Gowland P, Jones N, Robinson R. Diagnostic accuracy of antenatal magnetic resonance imaging (MRI) to predict birth weight greater than 90th centile or less than 10th centile in the third trimester. *Arch Dis Child Fetal Neonatal Ed* 2014;99(Suppl 1):A94–5.

52 Maticot-Baptista D, Collin A, Martin A, Maillet R, Riethmuller D. Prevention of shoulder dystocia by an ultrasound selection at the beginning of labour of foetuses with large abdominal circumference [French]. *J Gynecol Obstet Biol Reprod (Paris)* 2007;36:42–9.

53 Mazouni C, Rouzier R, Ledu R, Heckenroth H, Guidicelli B, Gamerre M. Development and internal validation of a nomogram to predict macrosomia. *Ultrasound Obstet Gynecol* 2007;29:544–9.

54 McLaren RA, Puckett JL, Chauhan SP. Estimators of birth weight in pregnant women requiring insulin: a comparison of seven sonographic models. *Obstet Gynecol* 1995;85:565–9.

55 Melamed N, Yogev Y, Meizner I, Mashiach R, Ben-Haroush A. Sonographic prediction of fetal macrosomia: the consequences of false diagnosis. *J Ultrasound Med* 2010;29:225–30.

56 Miller JM Jr, Korndorffer FA III, Gabert HA. Fetal weight estimates in late pregnancy with emphasis on macrosomia. *J Clin Ultrasound* 1986;14:437–42.

57 Miller JM Jr, Korndorffer FA Jr, Kissling GE, Brown HL, Gabert HA. Recognition of the overgrown fetus: in utero ponderal indices. *Am J Perinatol* 1987;4:86–9.

58 Miller J, Brown HL, Khawli OF, Pastorek IIJ, Gabert HA. Ultrasonographic identification of the macrosomic fetus. *Am J Obstet Gynecol* 1988;159:1110–4.

59 Nahum GG, Pham KQ, McHugh JP. Ultrasonic prediction of term birth weight in Hispanic women. Accuracy in an outpatient clinic. *J Reprod Med* 2003;48:13–22.

60 Nahum GG, Stanislaw H. A computerized method for accurately predicting fetal macrosomia up to 11 weeks before delivery. *Eur J Obstet Gynecol Reprod Biol* 2007;133:148–56.

61 Nicod AC, Hohlfield P, Vial Y. Performance of ultrasound estimation of fetal weight in fetuses weighing <2000 g or >4000 g. *Rev Med Suisse* 2012;8:2022–4.

62 Parry S, Severs CP, Sehdev HM, Macones GA, White LM, Morgan MA. Ultrasonographic prediction of fetal macrosomia. Association with cesarean delivery. *J Reprod Med* 2000;45:17–22.

63 Pates JA, McIntire DD, Casey BM, Leveno KJ. Predicting macrosomia. *J Ultrasound Med* 2008;27:39–43.

64 Pedersen JF, Molsted-Pedersen L. Sonographic estimation of fetal weight in diabetic pregnancy. *BJOG* 1992;99:475–8.

65 Peregrine E, O'Brien P, Jauniaux E. Clinical and ultrasound estimation of birth weight prior to induction of labor at term. *Ultrasound Obstet Gynecol* 2007;29:304–9.

66 Petrikovsky BM, Oleschuk C, Lesser M, Gelertner N, Gross B. Prediction of fetal macrosomia using sonographically measured abdominal subcutaneous tissue thickness. *J Clin Ultrasound* 1997;25:378–82.

67 Pollack RN, Hauer-Pollack G, Divon MY. Macrosomia in postdates pregnancies: the accuracy of routine ultrasonographic screening. *Am J Obstet Gynecol* 1992;167:7–11.

68 Rosati P, Exacoustos C, Pieroni A, Puggioni GF, Mancuso S. [Ultrasonic study of the predictability of fetal macrosomia (Italian). *Minerva Ginecol* 1990;42:239–42.

69 Rotmensch S, Celentano C, Liberati M, Malinger G, Sadan O, Bellati U, et al. Screening efficacy of the subcutaneous tissue width/femur length ratio for fetal macrosomia in the non-diabetic pregnancy. *Ultrasound Obstet Gynecol* 1999; 13:340–4.

70 Salomon LJ, Bernard JP, Duyme M, Ville Y. Predicting late-onset growth abnormalities using growth velocity between trimesters. *J Matern Fetal Neonatal Med* 2005;17:193–7.

71 Santolaya-Forgas J, Meyer WJ, Gauthier DW, Kahn D. Intrapartum fetal subcutaneous tissue/femur length ratio: an ultrasonographic clue to fetal macrosomia. *Am J Obstet Gynecol* 1994;171:1072–5.

72 Sokol RJ, Chik L, Dombrowski MP, Zador IE. Correctly identifying the macrosomic fetus: improving ultrasonography-based prediction. *Am J Obstet Gynecol* 2000;182:1489–95.

73 Sood AK, Yancey M, Richards D. Prediction of fetal macrosomia using humeral soft tissue thickness. *Obstet Gynecol* 1995;85:937–40.

74 Sritippayawan S, Anansakunwat W, Suthantikorn C. The accuracy of gestation-adjusted projection method in estimating birth weight by sonographic fetal measurements in the third trimester. *J Med Assoc Thai* 2007;90:1058–67.

75 Stein W, Delfy A, Schmidt S. Prediction of shoulder dystocia—combining foetal weight estimation by ultrasound and maternal risk factors–a solution for the dilemma? [German]. *Z Geburtshilfe Neonatol* 2009;213:180–5.

76 Sylvestre G, Divon MY, Onyeije C, Fisher M. Diagnosis of macrosomia in the postdates population: combining sonographic estimates of fetal weight with glucose challenge testing. *J Matern Fetal Med* 2000;9:287–90.

77 Tamura RK, Sabbagha RE, Depp R, Dooley SL, Socol ML. Diabetic macrosomia: accuracy of third trimester ultrasound. *Obstet Gynecol* 1986;67:828–32.

78 Hassibi S, Farhataziz N, Zaretsky M, McIntire D, Twickler DM. Optimization of fetal weight estimates using MRI: comparison of acquisitions. *AJR Am J Roentgenol* 2004;183:487–92.

79 Garden AS, Weindling AM, Griffiths RD, Martin PA. Assessment of fetal well-being with magnetic resonance. *J Perinat Med* 1991;19:435–48.

80 Baker PN, Johnson IR, Gowland PA, Hykin J, Harvey PR, Freeman A, et al. Fetal weight estimation by echo-planar magnetic resonance imaging. *Lancet* 1994;343:644–5.

81 Perinatal Institute. Ultrasound standards estimating fetal weight. 2014. [www.pi.nhs.uk/ultrasound/standards/EFW.htm]. Accessed 11 March 2015.

82 Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body and femur measurements – a prospective study. *Am J Obstet Gynecol* 1985;155:333–7.
83 Riley RD, Deeks JJ. The interpretation of random-effects meta-analysis. *BMJ* 2011;342:d549.
84 Bennini JR, Marussi EF, Barini R, Faro C, Peralta CF. Birth-weight prediction by two- and three-dimensional ultrasound imaging. *Ultrasound Obstet Gynecol* 2010;35:426–33.
85 Lindell G, Marsal K. Sonographic fetal weight estimation in prolonged pregnancy: comparative study of two- and three-dimensional methods. *Ultrasound Obstet Gynecol* 2009;33:295–300.
86 Nardozza LM, Vieira MF, Araujo JE, Rolo LC, Moron AF. Prediction of birth weight using fetal thigh and upper-arm volumes by three-dimensional ultrasonography in a Brazilian population. *J Matern Fetal Neonatal Med* 2010;23:393–8.
87 Schild RL, Fimmers R, Hansmann M. Fetal weight estimation by three-dimensional ultrasound. *Ultrasound Obstet Gynecol* 2000;16: 445–52.
88 Srisantiroj N, Chanprapaph P, Komoltri C. Fractional thigh volume by three-dimensional ultrasonography for birth weight prediction. *J Med Assoc Thai* 2009;92:1580–5.
89 Yang F, Leung KY, Hou YW, Yuan Y, Tang MH. Birth-weight prediction using three-dimensional sonographic fractional thigh volume at term in a Chinese population. *Ultrasound Obstet Gynecol* 2011;38:425–33.
90 Hadlock FP, Harrist RB, Fearneyhough TC, Deter RL, Park SK, Rossavik IK. Use of femur length/abdominal circumference ratio in detecting the macrosomic fetus. *Radiology* 1985;154: 503–5.
91 Hadlock FP, Harrist RB, Carpenter RJ, Deter RL, Park SK. Sonographic estimation of fetal weight. The value of femur length in addition to head and body measurements. *Am J Obstet Gynecol* 1985;151:333–7.
92 Hadlock FP, Harrist RB, Carpenter RJ, Deter RL, Park SK. Sonographic estimation of fetal weight. *Radiology* 1984;150:535.
93 Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology* 1991;181:129–33.
94 Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Ann Intern Med* 2013;158:544–54.
95 Hansen A, Chauhan SP. Shoulder dystocia: definitions and incidence. *Semin Perinatol* 2014;38:184–8.