Fetal macrosomia and delivery outcomes

Rebecca Robinson

School of Medicine

Department of Obstetrics, Gynaecology and Child Health

Table of Contents

Table of Figures

Table of Tables

Abbreviations

AC	Abdominal Circumference
ACOG	American College of Obstetricians and Gynaecologists
AD	Abdominal Diameter
BMI	Body Mass Index
BPD	Biparietal Diameter
BPI	Brachial Plexus Injury
CD	Chest Diameter
CI	Confidence Interval
DNA	Deoxyribonucleic Acid
EFW	Estimated Fetal Weight
FL	Femur length
g	gram
GDM	Gestational Diabetes
HC	Head Circumference
kg	kilogram
IPD	Individual Patient Data
LGA	Large for Gestational Age
mm	millimetre
NIHR	National Institute of Health Research
PTSD	Post-Traumatic Stress Disorder
RCOG	Royal College of Obstetricians and Gynaecologists
RCT	Randomised Controlled Trial
SGA	Small for Gestational Age
UK	United Kingdom
US	Ultrasound scan
USA	United States of America

Acknowledgements

"Peace – it does not mean to be in a place where there is no noise, trouble or hard work. It means to be in the midst of those things and still be calm in your heart" – unknown

I wish to dedicate this thesis to my gentle, generous and loving Mum who saw me begin but never got to see me finish. This has been written during an incredibly difficult period of my life where my own illness has repeatedly halted progress and many changes have ensued for my three sons and I.

I would firstly like to thank Nia Jones, my supervisor, for her unending patience and understanding. Her wisdom, direction and endless encouragement has enabled me to reach the end. Also, thank you to my second supervisor George Bugg for his relaxed interest and ideas. Thank you to Jim Thornton for challenging my assumptions and testing my understanding through debate. I would like to thank, Victoria White, Kate Walker and Kym Snell for their help and input to the process of my research.

My greatest thanks go to my partner Kirk Main for continuously encouraging and supporting me to finish, whilst being the most amazing Dad to our five children to allow me to write. A final thank you goes to my Ockbrook 'Busy Beaver' friends, who repeatedly suggested I give up because of everything else I was handling, but in fact, they just made me more determined to finish.

1 Abstract

Background

Fetal macrosomia occurs in 10% of all pregnancies and is associated with an increased risk of maternal and neonatal morbidity. Antenatal detection of fetal macrosomia by ultrasound is notoriously poor. No consensus exists on the measurement definition of fetal macrosomia, therefore this study investigates the association between various tests that are used to predict fetal macrosomia and adverse maternal and neonatal outcomes.

Objectives

To determine by systematic review which ultrasound measurement for predicted fetal macrosomia most accurately predicts adverse delivery and neonatal outcomes.

Search strategy

Four biomedical databases (Medline, Embase, PubMed, and the Cochrane Library) were searched for trials and observational studies published after 1966.

Selection criteria

Randomised controlled trials (RCT) or observational studies of pregnant women, with singleton pregnancies, resulting in a term birth (≥37 weeks) who have undergone an index test of interest measured and recorded as predicted fetal macrosomia during the third trimester of pregnancy (≥28 weeks).

Data collection and analysis

Adverse delivery and neonatal outcomes of interest included shoulder dystocia, brachial plexus injury (BPI), caesarean section, neonatal trauma, obstetric anal sphincter injury, and postpartum haemorrhage.

Results

Twenty-five observational studies (13,285 participants) were included. For BPI, the only significant positive association was found for Abdominal Circumference (AC) to Head Circumference (HC) difference >50mm (OR 7.2, 95%CI 1.8 to 29.0). Shoulder dystocia was significantly associated with abdominal diameter (AD) minus biparietal diameter (BPD) ≥2.6cm (OR 5.27, 95%CI 2.85-9.77, PPV 10.8%) and AC >90th centile (OR 2.72, 95%CI 1.24 to 5.96, PPV 8.6%) but not with an estimated fetal weight (EFW) >4000g (OR 1.97 95%CI 0.90-4.28, PPV 7.2%). Caesarean section was significantly associated with an EFW >4000g (OR 4.98, 95%CI 3.07-8.06, PPV 62%) and an EFW >90th centile (OR 2.90 95%CI 2.23-3.76)

Conclusions

Estimated fetal weight is the most widely used ultrasound marker to predict fetal macrosomia in the UK. However, this study suggests that other markers (AD-BPD and AC) have a higher positive predictive value for adverse outcomes associated with fetal macrosomia. My research suggests that we could be using better tests for fetal macrosomia and therefore further research is needed to instruct the RCTs for interventions aimed at reducing the adverse outcomes of fetal macrosomia.

2 Introduction

2.1 The purpose of identifying fetal macrosomia and the scope of the problem

Consensus has not yet been reached on the definition of antenatal fetal macrosomia. There are a multitude of imaging measurement definitions used to predict it. Many studies have been published testing the diagnostic accuracy of antenatal prediction of fetal macrosomia in comparison to absolute birth weights. It is also known, from a very large observational study, what level of absolute birth weight is associated with complications (1). These studies are useful for retrospective analysis but do not help to guide antenatal clinical management. What is not currently known is if any of these antenatal imaging measurement predictions are (more or less) associated with adverse maternal and neonatal outcomes. I have chosen to consider this and believe that by studying the association of the predicted fetal macrosomia ultrasound definition with adverse outcome, rather than the test accuracy of ultrasound for the comparison of predicted to actual birth weight, I could develop more clinically useful tools and statistics.

Good quality Randomised Controlled Trial (RCT) data has been published (and further studies are being undertaken) demonstrating that the incidence of adverse birth outcomes associated with macrosomia can be reduced by early induction of labour, without an unwanted increase in

the caesarean section rate (2-6). The ability of proven interventions to improve outcomes is an excellent step forward but the problem remains that we require a well-defined, accurate predictor of macrosomia that is associated with adverse outcome. Without this knowledge, we do not know who is best to target the interventions at, to yield the most cost effective and medically efficacious outcomes. An evidence-based definition is needed, especially now that we have a potential effective intervention.

We know that ultrasound prediction of fetal weight is mediocre in the case of big babies (7). Many described ultrasound measurement methods are individually more sensitive or specific but none currently combine both effectively. The ideal aim and best solution to this conundrum of detecting fetal macrosomia in the antenatal period would be a technique that is accurate and simple to implement. A proven, better prediction from ultrasound (US) may therefore be the current ultimate goal.

2.2 Definition of fetal macrosomia

Fetal macrosomia is the term used to describe a baby who is larger than average in utero or at the time of birth. Fetal macrosomia is defined in many ways. Using a statistical approach, any fetus weighing >90th centile for gestational age can be considered large for gestational age (LGA). Some researchers prefer to use the 95th centile as this corresponds to 1.9 standard deviations above the mean and therefore results in 90% of the population being defined as a normal weight. Some studies even use the 97.75th centile as this defines 95% of the population as being a

normal weight by corresponding to 1.96 standard deviations above the mean (8). The most commonly used 90th centile equates to a fetal birth weight of 4kg at 40 weeks gestation (9). Many institutions have therefore used 4kg as the definition, but currently, an estimated weight of >4.5kg is more widely used as a threshold to define fetal macrosomia (10-12). The American College of Obstetricians and Gynaecologists (13) recommends the use of a 4.5kg cut off, irrespective of the gestational age at birth. This is advised because babies that weigh more than 4.5kg have an increased risk of adverse maternal and perinatal outcomes (14, 15). Approximately 10% of all pregnancies result in fetal macrosomia but depending on the source and setting, this estimate ranges from approximately 3-15% (11, 16).

2.2.1 Aetiology

There are a number of risk factors associated with giving birth to a large birthweight infant, including maternal diabetes, both pre-existing and in pregnancy, a previous history of fetal macrosomia, maternal obesity, excessive maternal weight gain during the pregnancy, male fetus and multiparity (8, 17). Some of these risks, particularly diabetes, maternal weight gain and obesity, are associated with intermittent periods of hyperglycaemia in the mother and consequentially hyperglycaemia in the fetus via the placenta. Fetal hyperglycaemia results in stimulation of the production of insulin, insulin-like growth factor, growth hormone and other growth factors. These in turn lead to glycogen and fat deposition, causing excessive fetal growth. Another contributing factor is that

pregnancies that reach an advanced gestation will allow fetuses to grow in utero for longer and therefore potentially attain a greater birth weight.

There are a number of pathological causes for fetal macrosomia, although these are rare. These rare syndrome aetiologies should be considered when macrosomia is diagnosed in the presence of structural fetal anomalies and if the more common factors including obesity, diabetes and other constitutional factors have been excluded. Syndromes that are associated with fetal overgrowth include Beckwith-Wiedermann, Pallister-Killian, Sotos, Perlman, Simpson-Golabi-Behmel, Weaver, Costello and Macrocephaly Cutis Marmorata Telangiectasia Congenita.

2.2.2 Current thinking in assessment of fetal growth

There are two clear schools of thought advocated in the field of fetal and neonatal growth assessment (18). One is based on the idea that that fetal growth is strongly influenced by genetics and phenotypic factors such as the maternal height, weight and ethnicity (19). According to this theory, babies of different nationalities and populations are naturally expected to be different weights and sizes. Traditionally, fetal growth evaluation has been assessed by comparing the estimated fetal weight (EFW) to the population-based norms. These population-based standards are derived from either highly selected patient or heterogeneous cohorts, and by their nature, therefore, are a form of customised standards (20, 21). An inherent problem with these population-based standards is that they do

not differentiate between the naturally constitutionally large, but otherwise healthy fetus, and the fetus with abnormal macrosomic growth. When population based standards are used this can therefore result in under or over diagnosis of fetal macrosomia. In order to try to counteract the limitations of these population-based standards, customised normal growth charts were then created. These customized methods attempt to model ideal fetal growth potential by accounting for individual variables that are known to affect growth as opposed to assessing the deviation of growth from the expected population normal. One of the most commonly and widely used models is from Gardosi et al. (14, 22). Gardosi's group recommend using >90th customised GROW centile as the definition of LGA, based on a number of studies that have shown that customised growth centiles can improve the identification of macrosomia that is associated with pathological outcome (23-25). However, there are more recent reports that refute their conclusions (26, 27). The GROW technology was particularly targeted towards identification of the small for gestational age (SGA) fetus to reduce stillbirth rates. Customisation however is a contentious principle. For factors unrelated to the pathology suspected, customisation is appropriate but for factors that lie on the pathway, which leads to pathology, customisation is inappropriate (for example, smoking is a factor on the pathway to stillbirth and hence would be an inappropriate customisation factor). Customisation in these inappropriate circumstances would lead to the under diagnosis of the pathology sought.

In 2014, a different approach was recommended in published research. The Intergrowth-21st project (9) said that there is an ideal growth profile for all fetuses and newborns and that the observed global variance depends on the health, diet, environmental and lifestyle differences between the different populations. The results of the Intergrowth-21st project support the idea that poor growth is not an inevitable consequence of living in poor countries and that with the right support and investment into improving maternal and neonatal health, this inequality could change for the positive. In the INTERGROWTH-21st project, they excluded women with obstetric complications and conditions that may influence fetal growth and they saw less variability in fetal growth between the eight included countries (Brazil, Italy, Oman, UK, USA, China, India and Kenya).

The aim of the Intergrowth-21st project was to create new global fetal growth standards but it also intended to improve the monitoring and recognition of poor fetal growth and development. By the study's very nature though, the findings are just as applicable in settings where women and children are over-nourished. In many parts of the world, high-calorie but low nutrient, foods are becoming more and more accessible and common place in peoples' diets. This has led to higher rates of women being overweight, obese women becoming pregnant and consequently the higher incidence of gestational diabetes (GDM) and preeclampsia (28, 29). Advocates of this model feel that prior to this publication, the field of clinical research into fetal growth had been incredibly heterogeneous, with no consensus on what normal growth actually is. Ultimately, the creating

of this universally applicable growth standard tool has allowed the foundation of more homogeneous and generalizable research findings.

A second group of World Health Organisation (WHO) researchers were completing a very similar study to Intergrowth-21st and published in 2017 (30). Both of these studies sought universal fetal growth standards derived from healthy, normal mothers having healthy, normal pregnancies and concealed the ultrasound measurements. Intergrowth-21st and the WHO studies created the only two charts that have used modern techniques, without bias and had large enough samples to be precise. The WHO study also used multiple countries and weights and showed that fetal weights were affected by maternal age, height, parity and fetal sex. The researchers recognised that none of these effects was large yet published separate estimated fetal weight charts by country. Despite this, the final charts from the WHO study were similar to those of Intergrowth-21st. There were three main differences between WHO and Intergrowth-21st. Firstly, WHO was smaller in sample size than Intergrowth (1387 vs 4321 patients). Secondly, the Intergrowth-21st study did not reveal the scan measurements on the screen as the callipers were placed, whilst WHO did reveal them, which may have introduced bias. Thirdly, WHO used Hadlock's formula to estimate EFW which is a popular but outdated formula that was based on a small number of women (21). Intergrowth-21st took into account that charts based on babies that are actually born preterm, underestimate centiles as babies that are born pre-term are commonly growth restricted. They did this by creating their own new formulae for estimating fetal weight that was more evidence based and had less bias

due to their larger sample and more modern scanning/weighing techniques. Due to these factors, the Intergrowth-21st charts are marginally preferred and used in some institutions.

To date there has not been a clinical trial that directly compares the two methods to adverse fetal or neonatal outcomes and thus we are inherently reliant on data from observational studies. A prospective cohort study by Pasupathy et al. demonstrated that customised centiles are more strongly associated with adverse birth outcomes in LGA neonates (25). Unfortunately, this trial compared the customised centiles with population centiles that are themselves a form of customised chart. Empirical comparisons such as this one by Pasupathy are therefore actually not helpful. What was needed was to compare the different chart's detection rates for adverse fetal outcomes. When the detection rate for a pathology improves, it is common for the false positive rate also to be higher. In this scenario, detecting more babies with adverse outcomes is not necessarily better, if the intervention for the detection of these babies has complications of its own. It is therefore important that a good empirical study, to compare the different chart's detection rates, fixes the false positive rate.

There are at least three studies that have suggested that customised charts could have detected more pathology than Intergrowth-21st but they all found that using the customisation charts created double the false

positive rates (31-33). A recent, large US study by Kabiri (34) compared five different charts that included Intergrowth, WHO and GROW. This good quality study compared different adverse outcomes for a fixed false positive rate of 10%. They showed that intergrowth had the highest sensitivity and GROW the lowest, although the confidence intervals were overlapping. This study was biased by treatment paradox as the scan findings were known by the obstetricians and therefore their clinical decision-making may have influenced the adverse outcomes. Only one study has managed to avoid this treatment paradox by concealing the ultrasound scan findings from the clinicians (35). The Pregnancy Outcome Prediction Study (POPs) study, a prospective cohort from Cambridge, showed that customisation using GROW did not improve the detection rates of any adverse outcomes.

2.2.3 Global variance

There is a worldwide epidemic of obesity and obesity related diseases such as diabetes. In certain nations such as the United States obesity rates have reached an all-time high (36). In the United Kingdom (UK) today we have a diverse, multi ethnic population and the rates of fetal macrosomia (approximately 10% (37)) reflect both this and an increasingly obese, less physically active population. Maternal obesity and diabetes are independent risk factors for fetal macrosomia (11). The prevalence of macrosomia in developed countries is between 5% and 20%; however, an increase of 15–25% has been reported in the past two to three decades, mainly driven by an increase in maternal obesity and

diabetes (38). The prevalence of macrosomia (defined as birth weight ≥4000g) in developing countries is typically 1-5% (with a range from 0.5-14.9%) (39). There is scare data available for the changing prevalence of macrosomia in developing countries but investigators of one study in China (40) noticed an increase from 6% in 1994 to 7.8% in 2005. There is an increasing prevalence of diabetes and obesity in women of reproductive age in developing countries (41, 42) and it might therefore be expected that a parallel increase in macrosomic births may follow. The overall, worldwide prevalence of infants ≥4000 g at birth is approximately 9% and approximately 0.1% for weight ≥5000 g, with wide variations among countries (7). The lowest prevalence of macrosomia found was ≤3% in reports from Nigeria (43), Pakistan (44), and Thailand (45). The highest prevalence of ≥20% was reported in Taiwan (46), Denmark (47) and the Republic of Croatia (48). When considering a worldwide definition of fetal macrosomia, many studies, including the WHO surveys (39), have described fetal macrosomia as being present when birthweights are greater than the 90th centile of the country-specific birthweights. For example, this equates to a definition of 4050g in Algeria but just 3250g in India. It is debated how much of this global variation is due to ethnicity and it's associated phenotypic factors, when advocates of the Intergrowth-21st publication think that health, diet, lifestyle and environmental factors underlie these differences in fetal weight. The conclusions of the Intergrowth-21st study would tend to disagree with this country-specific definition, as they suggest that there is a normal growth profile for all fetuses worldwide. A definition of macrosomia of ≥4000g correlates with a

newborn weight that plots between the 90th-95th centile on the Intergrowth-21st International Standards Chart (49).

Globalisation policies have increased the rates of human migration from developing to developed nations. These migrant populations typically adopt an urban, western lifestyle. Koyanagi and colleagues published a study on behalf of the World Health Organisation (WHO) that described a survey of fetal macrosomia in 23 different developing countries, from the three continents of Latin America, Africa and Asia (39). Their findings showed that both maternal BMI and diabetes were strong risk factors for macrosomia. Mothers of babies with macrosomia were at risk of caesarean section because of obstructed labour and post-term pregnancy, and for adverse maternal outcomes in all regions. In addition, macrosomia was associated with an increased risk of adverse perinatal outcomes only in Africa. This study only included institutions that had over 1000 deliveries per year and where caesarean section was performed. There was therefore unavoidable selection bias, as in most developing countries the overall majority of deliveries occur in a home birth setting. The Lancet 2013 publication by Dennedy (50) and colleagues makes a broad comparison of this WHO study with other Demographic and Health Surveys (51), which selected births from both rural and urban cohorts across Sub-Saharan Africa. The comparison publication (50) reports consistently increased rates of obesity found in the mothers of the countries that were included in both studies. It is possible that additional bias will have influenced the maternal and perinatal indices, as a result of the urban demographic of the studies. However, in consideration of this limitation, they found that, like

the already existing evidence for the western world (38), there is also an increasing prevalence of gestational obesity and diabetes in developing countries, that are rapidly becoming more urban and adopting western cultures. Consequentially, this could be associated with a parallel increase in fetal macrosomia in developing countries.

The data from the studies in the developing world also support a body of evidence (52) that recommends an active approach to the early identification and management of pre-gestational and gestational diabetes in women from migrant and ethnic populations when they first present to primary care and to antenatal clinics. As the health care and funding for assistance is limited in developing countries, they often rely on aid from the developed world. Therefore, the health care assistance provided currently tends to be focused on the higher priority issues such as HIV and maternal mortality (50). Macrosomia is associated with the long term risks of metabolic disorders and obesity (53-58), so inaction in the developing world regarding fetal macrosomia will inevitably result in future health problems for these populations.

2.3 Impact of fetal macrosomia and adverse outcomes

Maternal and neonatal morbidity is greater when the infant is macrosomic. Due to their size and body fat distribution, these babies and their mothers are at increased risk of complications at the time of birth,

particularly shoulder dystocia, when the baby's shoulders are stuck at the level of the bones of the maternal pelvis. These complications are associated with adverse maternal and neonatal outcomes. For the mother these include prolonged labour, an increased caesarean section rate, postpartum haemorrhage, more severe perineal tears including third and fourth-degree tears (59, 60). The macrosomic neonate is at risk of shoulder dystocia, fetal asphyxia, birth trauma and neonatal hypoglycaemia (61). These outcomes not only impact on the health of mothers and their newborn infants, they can also result in significant costs to the NHS as a result of the consequential long term care requirements and occasionally the settlement of litigation cases if negligence is proven (62). In addition to this, there are recognised associations between fetal macrosomia and more long-term consequences including diabetes, heart disease and obesity [43-48].

The estimation of the level of these risks was illustrated well by a very large study that included all singleton live births in the United States from 1995 to 1997 (1). They assessed birth weight categories to determine predictive thresholds of adverse outcomes. The research used logistic regression analysis to compare macrosomic infants within three weight criterion groups (4000-4499grams, 4500-4999grams and >5000grams) to normosomic infants (3000-3999grams). Compared to appropriate-for-gestational-age infants the relative risk of birth injuries was twofold, threefold and fourfold greater for infants with birth weights of grade one (4000-4499g), two (4500-4999g) and three (>5000g), respectively. Their results showed that the frequency of complications increased with the

severity of the macrosomia. They concluded that although a definition of macrosomia as >4000g may be useful for identifying increased risks of labour and newborn complications, >4500g was more predictive of neonatal morbidity and >5000g a better indicator of infant mortality risk. Therefore, the risk of the severity of the adverse outcome increases along a continuum based on the degree of fetal macrosomia.

2.3.1 Fetal complications

Shoulder dystocia leading to birth trauma or long-term disability is the most common serious intrapartum concern. The incidence of shoulder dystocia varies from 0.58-0.70% of all deliveries (63) and 11% of macrosomic deliveries (64). In a large published population study (65) a multivariable analysis demonstrated an odds ratio of 16.1 (CI 13.2-19.6, P<0.001) for the risk of shoulder dystocia when the birth weight was ≥4kg. The risk of injury or adverse outcome is positively associated with increasing fetal weight (1). The RCOG defines shoulder dystocia as 'A vaginal cephalic delivery that requires additional obstetric manoeuvres to deliver the fetus after the head has delivered and gentle traction has failed' (66). An objective diagnosis of a prolongation of head-to-body delivery time of more than 60 seconds has also been proposed (67) (68) but these data are not routinely collected. Shoulder dystocia occurs when after the delivery of either the fetal head, the anterior, or less commonly the posterior, fetal shoulder impacts on the maternal symphysis, or sacral promontory, respectively. Although this risk of shoulder dystocia increases with the weight and size of the baby, 40-60% of shoulder dystocia cases

occur in infants that are not macrosomic or over 4000g (69). This makes the prediction of shoulder dystocia difficult and it therefore remains one of the most feared emergencies that obstetricians encounter in their practice. There are manoeuvres that have been described to help to aid delivery of the shoulders when shoulder dystocia occurs. There is a recommended and widely used algorithm that is taught in obstetric practice that employs the use of these manoeuvres in a specified order (63). This enables the management of this emergency to be conducted in a systematic way.

A clear association exists between macrosomia and shoulder dystocia. The risk factors for both of these overlap significantly, with women who are diabetic, obese or have had previous macrosomic babies being the most at risk (70). Studies have suggested that fetal macrosomia is an important independent risk factor for shoulder dystocia (61, 71). The most common fetal complication is brachial plexus injury (BPI). Recorded incidence ranges from 2-16% within the literature (72) but less than 10% of cases of BPI result in any long term disability or permanent injury (73, 74). Fractures of the humerus or clavicle can occur in approximately 10% of shoulder dystocia cases but these usually recover without any adverse consequences. Despite this clear association between macrosomia and shoulder dystocia, it is not proven useful as a predictive tool. Risk assessment models have been attempted by looking individually at, and combining known risk factors for shoulder dystocia but these methods have not demonstrated a good positive predictive value (75). This is in part because it is difficult to determine an accurate fetal size before birth but also because the vast majority of macrosomic infants (>4.5kg) do not

experience shoulder dystocia. In addition to this, almost 50% of shoulder dystocia cases occur when infants weigh less than 4kg. The RCOG Green Top Guidance (63) states "Risk assessments for the prediction of shoulder dystocia are insufficiently predictive to allow prevention of the large majority of cases". Therefore, shoulder dystocia prediction modelling is not recommended.

Although very rare, hypoxic-ischaemic brain injury can occur following shoulder dystocia. The literature documents the incidence of this occurring as ranging from 0.5-23% of cases of shoulder dystocia. The critical issue regarding the risk of hypoxia is the duration of time from delivery of the head to delivery of the body. The risk is very low with a duration of less than 5 minutes but once 5 minutes is exceeded the risk increases (76). Cases of perinatal mortality directly attributable to shoulder dystocia are incredibly rare (reported in 0.4% of cases) and these tend occur when the manoeuvres to release the trapped shoulder are not successful.

2.3.2 Maternal Complications

The exact incidence of maternal morbidity following the delivery of an infant with fetal macrosomia is not available. However, it is clearly the case that the specific risks associated with birth of a macrosomic fetus include protracted or arrested labour, operative vaginal delivery, caesarean delivery, genital tract lacerations, postpartum haemorrhage and rarely uterine rupture (12, 77, 78).

It is possible that fetal macrosomia is a greater risk still for women in developing countries. Cephalo-pelvic disproportion occurs more frequently because of undernutrition during youth and that women are more likely to become pregnant at a young age, before pelvic growth is complete (79). This is compounded by the fact that facilities for operative delivery in women with obstructed labour are not consistently available.

2.3.2.1 Genital tract lacerations

Fetal macrosomia and shoulder dystocia are associated with both precipitous and protracted labours. Both of these circumstances combined with the increased need for assisted delivery contribute to the increased risk of genital tract trauma. These genital tract lacerations can vary in severity from the simpler second-degree tears to the more serious third and fourth degree tears that involve disruption of the concentric muscles that surround the anus and preserve faecal continence. Severe trauma can also result in an increased rate of postpartum haemorrhage.

2.3.2.2 Caesarean Section

The ACOG bulletin emphasizes that an increased risk of caesarean delivery is the primary maternal risk factor associated with macrosomia (11). They report that cohort studies demonstrate that the incidence of caesarean section delivery in women who attempt a vaginal delivery is at least doubled when the fetal weight is estimated to be greater than 4500g (13). They go on to recommend that with an EFW of more than 4500g, a prolonged second stage of labour or an arrest of fetal descent in the second

stage should prompt caesarean delivery. This recommendation is quoted as level B evidence (based on limited or inconsistent scientific evidence) but has not been translated into routine UK practice. Published cost effectiveness reports (80) do not support prophylactic caesarean delivery for suspected fetal macrosomia if the EFW is less than 5000g (in the non-diabetic maternal population), although some authors agree that caesarean delivery in these situations should be considered.

2.3.2.3 Postpartum haemorrhage

The literature suggests that significantly more LGA pregnancies are complicated by postpartum haemorrhage (81, 82). A large retrospective study suggested that the risk of postpartum haemorrhage is statistically increased when birth weight is greater than 4500 g (OR 5.23, CI 1.63–16.80, $P<0.006$) but not for birth weight between 4000–4500 g (83). Postpartum haemorrhage that complicates fetal macrosomic birth is commonly due to uterine atony and genital lacerations, which can be managed effectively if guidelines are strictly followed (84-86).

2.3.3 Neonatal complications

In addition to trauma at birth, during the neonatal period, macrosomic infants are predisposed to electrolyte and metabolic disturbances, such as hypoglycaemia and hyperbilirubinemia. Many of the metabolic disturbances, such as hypomagnesaemia and hypocalcaemia occur more regularly in the offspring of diabetic mothers, so this will be reviewed in more detail during my discussion of diabetes.

2.3.3.1 Hypoglycaemia

Hypoglycaemia is defined in neonates as a blood glucose <40 mg/dL (2.2 mmol/L). Neonates that are born macrosomic have been subject to intrauterine exposure of excessive glucose, which results in hyperinsulinemia, increased utilization of oxygen and glucose, and oxidative stress. They are therefore used to a level of relative hyperglycaemia before birth. Once born, the placental supply of glucose is interrupted and the infant can develop hypoglycaemia. This incidence of hypoglycaemia in macrosomia infants is documented as ranging from 15-19% (87). The incidence was highest in infants of mothers with diabetes.

It can be asymptomatic but in the symptomatic infant the signs and symptoms reflect how the nervous system responds to glucose deprivation. These can include autonomic symptoms such as jitteriness, tremors, pallor, tachypnea and irritability. More severe neuroglyocopaenic symptoms can occur comprising poor feeding, lethargy, hypotonia, coma and seizures. Seizures are rare, occurring in up to 0.3% of LGA infants (87).

2.3.3.2 Respiratory Distress

LGA infants are more likely to develop respiratory distress than their normal sized counterparts (1). This most commonly occurs due to respiratory distress syndrome and the incidence is highest in the macrosomic infants of diabetic women, who tend to be delivered more prematurely. Transient tachypnoea of the newborn is also more frequent

but may be in part be due to the higher rate of caesarean section required for macrosomic infants.

2.3.3.3 Perinatal asphyxia

There is indirect evidence that macrosomic infants are at increased risk of perinatal asphyxia (88). LGA infants have a higher frequency of low Apgar scores than appropriate for gestational age infants. In addition to the risk of asphyxia from shoulder dystocia, it is likely that there is also some contribution from the increased oxidative stress the fetus experiences in utero due to hyperglycaemia and hyperinsulinemia.

2.3.4 Lifetime risks

In future life, macrosomic infants that are at the highest end of the distribution for weight or body mass index (BMI) are more likely than other infants to become obese in childhood, adolescence, and early adulthood (53, 54). They also are at risk of cardiovascular, diabetes and other metabolic complications in adult life (55-57).

Adults that were macrosomic at birth have impaired glucose tolerance and (particularly in the offspring of diabetic mothers) have reduced insulin secretion in later life (89). Infants born macrosomic are prone to developing cardiometabolic disease and the underlying mechanisms for this continue to be investigated. There is research to suggest that altered DNA methylation in neonates born LGA is associated with cardiometabolic risk in children (90).

2.4 Diabetic pregnancy and fetal macrosomia

2.4.1 Aetiology and differences in risks to the non-diabetic population

There are three different types of diabetes found in pregnancy. Pre-existing type 1 diabetes, pre-existing type 2 diabetes and GDM. The most common aetiology encountered in clinical practice is GDM. It is estimated that approximately 87.5% have GDM (which may or may not resolve after pregnancy), 7.5% have type 1 diabetes and the remaining 5% have type 2 diabetes (91). The diabetes can pre-exist the pregnancy and most commonly, this will be in the form of type 1, insulin dependent diabetes. The highest complication rates are found in the mothers with pre-existing diabetes, especially if their diabetic control is poor at conception and in the first trimester of pregnancy. High levels of maternal hyperglycaemia in the pre-conception period and early pregnancy can result in miscarriage and fetal anomalies (92, 93). There is an increasing prevalence of type 2 diabetes in mothers due to increasing obesity rates in the population; clinicians are therefore encountering this aetiology more frequently. It is thought that the risks associated with pre-existing type 2 diabetes are comparable to the severity of risks with pre-existing type 1 diabetes (94).

Infants of diabetic mothers who are born macrosomic are more likely than their non macrosomic counterparts to have shoulder dystocia, brachial plexus injury, respiratory distress, hyperbilirubinemia, hypoglycaemia and acidosis (95). The increased risk of respiratory distress syndrome is true for any given gestational age. This occurs because

maternal hyperglycaemia appears to delay surfactant production. It is thought that the underlying mechanism of this is that neonatal hyperinsulinaemia can interfere with the induction of lung maturation by glucocorticoids (96). The macrosomic infants can appear large and plethoric, with visceromegaly and have excessive fat deposition apparent in the abdominal and scapular areas. The shoulder dystocia risk in the macrosomic infants of diabetic mothers increases by 2 to 4 fold compared to infants of the same birth weight and length born to mothers without diabetes (61, 70).

It is this increased risk of shoulder dystocia in diabetic mother's babies compared to non-diabetic mothers having infants of the same weight that prompts the directions and advice of the Diabetes in Pregnancy guidelines that are used clinically worldwide (13, 91). Mothers with diabetes who are carrying fetuses predicted to be macrosomic are recommended elective caesarean section if their EFW is greater than 4.5kg. This contrasts with the recommendation of elective caesarean for predicted macrosomia in the non-diabetic population, which is only recommended if the EFW is greater than 5kg.

2.4.2 Pre-existing diabetes

Macrosomia can occur in all diabetic pregnancies but it seems to be most common in women who have pre-gestational diabetes. A large Swedish birth registry study over 10 years showed the incidence of LGA

infants as 47% in pre-existing diabetic pregnancies compared to 26% in GDM and 11% in the non-diabetic population (97, 98).

Higher risks of the complications of fetal macrosomia are reported in the infants of pre-gestational diabetic mothers. The Confidential Enquiry into Maternal And Child Health from 2002 and 2003 demonstrated a number of these risks (99). They reported a doubling of the risk of fetal macrosomia (22% versus 11%) and an almost three times risk of shoulder dystocia (8% versus 3%). These risks were generated by comparing the women with pre-gestational diabetes to the general maternity population. This increased risk of shoulder dystocia resulted in a ten fold increased risk of brachial plexus injury (100).

2.4.3 Gestational diabetes

Despite being less common than in the pre-existing diabetic population, the offspring of gestational diabetic pregnancies also have a higher incidence of fetal macrosomia. Macrosomic infants grow disproportionately; they have an increased ponderal index that results in an increased ratio of shoulder and chest to head measurement. These infants have higher body fat and thicker skin folds in their upper limbs compared with infants of similar weights who were the offspring of non-diabetic mothers (97, 101). There is evidence that mothers with gestational diabetes are more likely to grow fetuses with disproportionate macrosomia than the LGA infants of nondiabetic mothers (98).

2.4.4 Prediction of fetal macrosomia in diabetic pregnancy

There is still no clear consensus for the best way to monitor fetal size in pregnant women with diabetes (91). Most commonly, sonography is used and the weight is predicted using the abdominal circumference (AC) and EFW of the fetus. The formulae used for calculating EFW are less accurate in women with diabetes and in macrosomic babies, compared to their non-diabetic and normosomic counterparts (102, 103). The babies of women with diabetes who are LGA have an increased growth of tissue that is insulin sensitive, such as abdominal wall fat. It is for this reason that sonographic fetal AC is widely considered the most clinically useful measure of diabetes related macrosomia and shoulder dystocia. AC is also the most accessible measure in later pregnancy when the fetal head becomes engaged within the pelvis and becomes too low to measure accurately (91).

2.4.5 Prevention of macrosomia in diabetes

In the management of women with diabetes in pregnancy it is highly recommended and important that pregnancy is planned and that conception occurs when glycaemic control in pregnancy is optimal (91). The better the glycaemic control is, the less risk there is of fetal macrosomia developing. Pregnant women must be advised to limit their weight gain during pregnancy through a healthy diet, lifestyle and exercise. This advice is the same to all types of pregnancy, with or without diabetes. Type 2 diabetic pregnant women should be left on metformin but

many will require the transition to insulin injections to help control excessive fetal growth and their own varying glycaemia (104).

2.5 Antenatal prediction of fetal macrosomia

Antenatal prediction of fetal macrosomia is often inaccurate. Any strategy used to detect fetal macrosomia is therefore limited by the inaccuracy of the methods of fetal weight estimation.

2.5.1 Maternal

Maternal estimation of fetal weight has been shown in one study to be at least as accurate as clinical estimation, especially in parous women who have had a previous LGA infant (105). Although this method is not usually used clinically, it can be a useful adjunct in women who present in labour without any recently documented antenatal care or size estimation.

2.5.2 Clinical

The fetal size can be estimated by palpation and measurement of the maternal abdomen. Leopold's manoeuvre is a stepwise method of abdominal palpation that assesses and establishes the fetal lie, presentation and presenting part. This combined with the measurement of the fundal height can estimate the fetal size. The measurement is taken with a tape measure from the highest fundal endpoint (which is an operator dependent subjective measure) to the superior border of the symphysis pubis. Clinical estimation of fetal macrosomia is unreliable (106-108)

although serial measurements by the same practitioner over time can yield improved sensitivities. The prospectively measured sensitivity of clinical methods for detecting macrosomia are poor and range from ten to 43 percent meaning that a lot of macrosomic babies are not predicted before birth, however the specificity has been found to be higher than the sensitivity when macrosomia is predicted (99%) (109, 110).

2.5.3 Two-dimensional ultrasound

Ultrasound sonography is the most widely used and common modality for the diagnosis and prediction of fetal macrosomia and LGA. Evidence has shown that the accuracy of ultrasound in estimating the weight of large babies is approximately 45% sensitivity for EFW>4000g (111). A large variety of ultrasound measurements have been, and continue to be, used to estimate fetal weight and specifically predict macrosomia.

A number of factors limit the accuracy of sonography. These include the inaccuracy of the biometry measurements (although these can be reduced by repeated measures and careful checking of the anatomical sections (112)), inexperienced operators, the quality of the sonography equipment, oligohydramnios, maternal obesity and the inaccuracy of the formulae for EFW used.

2.5.3.1 Abdominal circumference

AC is widely reported as the most important biometric parameter for the identification of fetal macrosomia. This tends to be the most commonly used tool in decision-making and delivery planning in a clinical setting (113, 114). An AC of >35-38cm alone is predictive of fetal macrosomia but overall, the Coomarasamy systematic review and meta-analysis in 2005 concluded that there was no difference in accuracy between ultrasonographic EFW and AC in the prediction of a macrosomic baby at birth (115).

2.5.3.2 Estimated fetal weight

As the weight cannot be measured directly by ultrasound, biometric measures are combined into formulae to calculate an EFW. Most commonly, a combination of the head circumference (HC), biparietal diameter (BPD), abdominal circumference (AC) and femur length (FL) are used in a regression model to calculate an EFW.

The most popular formulae used for estimation of fetal weight are Hadlock 4 (there are also versions 1-3 described) and Shepard. They both use a combination of the biometric measurements previously described.

Hadlock 4 formula:

- Log10 Birth Weight = 0.3596+(0.00061 X BPD X AC) + (0.0424 X AC) + (0.174 X FL) + (0.0064 X HC) - (O.OO386 X AC X FL)

Shepard formula:

- Log10 Birth Weight = -1.7492 + 0.166 (BPD) + 0.046 (AC) -(2.646 [AC X BPD] /100)

The literature suggests that the Hadlock 4 formula of HC, AC, and FL, appears to perform best across a range of different populations and fetal weights and provides the most accurate estimation of fetal weight (116, 117), hence why this is the most commonly used measure.

There are almost 40 described formulae for the estimation of fetal weight but they tend to perform better for normally sized fetuses than LGA ones. The evidence suggests that no formula is clearly the best for macrosomia (118). A review investigating the prediction of macrosomia of >4000g showed widely varying sensitivities between the separate studies (12 to 75%) (7). The diagnosis of predicted macrosomia of >4500g is even less accurate (103). The literature suggests that there is only a 60% sensitivity for identifying macrosomia (of >4.5kg) antenatally and third trimester ultrasound scans have at least a 10% error margin when subsequently compared to actual birth weight (119, 120). There are studies that have tried to address this fact, that current formulae are not accurate for the prediction of macrosomia. Porter et al. concluded that their macrosomia-derived formula was significantly better at predicting macrosomia than the traditionally used Hadlock formula (121). However, they split their single dataset into two groups, of a training and a validation arm, to both derive and then test their formula. Therefore, it can only be

concluded that their macrosomia formula improves macrosomia detection within this patient sample. Further studies would be needed to assess the validity and efficacy of this formula in other settings and samples.

There is clearly controversy over the issue of fetal macrosomia prediction, as Coomarasamy and colleagues systematic review and meta-analysis (115) conclude that a positive US prediction is more accurate for ruling in macrosomia than a negative result is for ruling it out, whilst the ACOG Bulletin concludes the opposite (13).

2.5.3.3 Other techniques

2.5.3.3.1 Soft tissue measurements

The EFW and its component biometric parts do not take fetal body composition into account. As a result of this, the actual birth weight of babies with very similar biometric measurements can vary significantly. On average, body fat accounts for 14% of the birth weight, but 46% of birth weight variance (122). This has led to the development of different soft tissue measurements being attempted to predict macrosomia.

Soft tissue measurements have been recorded and data published for the shoulder (123), mid humerus (124, 125), thigh (126), abdominal wall (127) and cheek-to-cheek measurements (128). Prenatal sonographic adipose measurements have been reported to correlate well with postnatal

skin fold thicknesses but the data is limited because the sample populations are small (129). However, a study that compared soft tissue measurements to clinical and sonographic EFW did not indicate that soft tissue measurements were superior in identifying fetal macrosomia (107).

2.5.3.3.2 Composite measurements

Some studies have suggested and tested ratios or combinations of 2D US measurements in the prediction of fetal macrosomia and shoulder dystocia. Ratios investigated include HC/AC, AC-HC, Abdominal Diameter (AD)-BPD and Chest Diameter (CD) - BPD (130-135).

2.5.4 Three-dimensional ultrasound

There is limited evidence on the use and accuracy of Three-dimensional ultrasound especially with specific reference to the prediction of fetal macrosomia. Three-dimensional ultrasound has a better ability to assess fetal soft tissues and fetal volume but studies that have attempted to validate the method of three-dimensional ultrasound for predicting fetal weight have demonstrated that three-dimensional ultrasound was not superior to two-dimensional ultrasound fetal weight estimation (136, 137). However, two studies have found increased benefit from three-dimensional ultrasound. The first study (138) compared two-dimensional and three-dimensional ultrasound for general fetal weight estimation. They found that by combining two-dimensional biometry with fetal thigh volume, a greater proportion of EFWs were within 5% of the actual birthweight when this was compared to two-dimensional use of the Hadlock formula alone

(70% versus 40%). A more recent study assessed the accuracy of three-dimensional ultrasound fractional limb volume in comparison to two-dimensional ultrasound within a population of women with GDM. The three-dimensional method demonstrated better sensitivity than two-dimensional ultrasound for the prediction of fetal macrosomia (84% versus 63%) (139). However, other studies did not find any significant advantage to three-dimensional methods (136, 140, 141). It has been suggested that the best method for predicting macrosomia may be found by combining three-dimensional fetal volume measurements with two-dimensional biometry (142). They demonstrated that by combining these measurements, the percentage error was reduced to 6.5% compared to 10-15% with two-dimensional ultrasound alone. However, this study only had a small number of macrosomic infants, with a higher percentage error identified, so this technique may not perform as well for macrosomic compared to normal infants.

The only systematic review and meta-analysis published on this topic concluded that only one study reported the accuracy of three-dimensional ultrasound to specifically predict birthweight of >4000g. This demonstrated a sensitivity of 0.42 (95%CI 0.26-0.61) and a specificity of 0.98 (95% CI 0.95-1.00)(143). However, conclusions were unable to be drawn from one study alone (144).

2.5.5 Magnetic Resonance Imaging (MRI)

Published research suggests that the sensitivity and specificity of MRI is better than US for the diagnosis of macrosomia (144) but the studies are small and MRI is not widely available. MRI provides a better assessment of fetal fat than US so in theory this should be a better method of predicting fetal macrosomia (145) but MRI is not as widely available as sonography and is more expensive. A systematic review and meta-analysis (144) compared two-dimensional ultrasound with MRI for predicting a birth weight of >90th centile or >4000g. Malin and colleagues found that the MRI EFW was more sensitive to detect fetal macrosomia than the two-dimensional ultrasound EFW (0.93 to 0.56). This enabled a conclusion that MRI is more sensitive than US for detecting fetal macrosomia using EFW. The MRI EFW was not however statically significantly more sensitive than a two-dimensional ultrasound abdominal circumference of >35cm for detecting fetal macrosomia (0.93 versus 0.80), although it was more specific (0.95 to 0.86). They concluded that more primary research is required into this topic and for now MRI use should be confined to within research studies.

2.6 Intervention in fetal macrosomia

2.6.1 Initial prevention

The main risk factors for macrosomia are obesity, maternal weight gain and poor glycaemic control (146). The prevalence of macrosomia is increasing because the prevalence of maternal obesity is increasing worldwide, in both developed and developing countries. Approximately one

fifth of UK pregnant women and up to one third of USA pregnancies are obese (36). I have previously discussed that obesity in pregnancy increases the risks of both short and long-term adverse outcomes for mother and baby. Therefore, a mainstay of prevention of macrosomia must be in the prevention of obesity before pregnancy and the management of obesity in pregnancy.

It had been hypothesised that metformin may help to reduce the risk of fetal macrosomia due to its association with less gestational weight gain and the effect it has on reducing insulin resistance. However, recent clinical trials have shown that in non-diabetic obese mothers (BMI>30 and 35kg/m^2), a daily dose of metformin from 12-16 weeks until delivery did not result in any average reduction in neonatal rates of macrosomia or LGA infants (147).

The recent Lancet publication of the UK-UPBEAT (Pregnancies Better Eating and Activity Trial) RCTl did not demonstrate a statistically significant improvement in the incidence of macrosomia, LGA infants in women with obesity who made lifestyle changes during pregnancy (148). That said, it is fundamentally important that pre-pregnancy counselling and public health initiatives should emphasize the importance of attaining a healthy weight prior to pregnancy and also to avoiding excessive weight gain during pregnancy (146).

2.6.2 Induction for fetal macrosomia to reduce adverse outcome

A well designed, RCT was published in 2015 (149) that concluded that induction of labour for suspected LGA fetuses is associated with a reduced risk of shoulder dystocia and associated morbidity compared with expectant management. In this study an EFW of >95th centile was used. It was found that induction of labour did not increase the risk of caesarean delivery and in fact improved the likelihood of normal vaginal delivery. The reader was advised that these benefits should be balanced against the effects of early-term induction of labour.

A Cochrane library systematic review was published in 2016 (2) that included four randomised studies (including the afore mentioned RCT). Many doctors have changed their practice and feel the evidence already justifies induction but others disagree. The authors concluded that further trials of induction shortly before term for suspected fetal macrosomia remain justified. Such trials should concentrate on refining the optimum gestation of induction and improving the diagnosis of macrosomia. This research area is highly relevant and topical. The Health Technology Assessment Programme of the National Institute of Health Research (NIHR) has identified it as immediately useful to patients, clinical practice and policy/decision makers. They have gone on to fund a trial of induction of labour for predicted macrosomia, the 'Big Baby Trial' (150), which is investigating whether a policy of induction of labour at 38+ weeks gestation in women with babies considered LGA (defined as >90th

customised GROW centile of EFW) will reduce the incidence of shoulder dystocia. They aim to finish recruitment in 2021.

2.7 Summary of current evidence for the association of predicted macrosomia and adverse outcomes

There is a paucity of published literature linking predicted fetal macrosomia with outcome. There are no diagnostic accuracy studies published that examine the estimation of fetal macrosomia by ultrasound being associated with or predictive of adverse outcomes. The background search results consist of observational studies, which are predominantly using retrospective cohorts in their study design. There is a large volume of this observational data published that has recorded the correlation and links between the absolute birth weights of macrosomic infants and the adverse outcomes that they are more likely to encounter. Similarly, many published articles have studied the diagnostic accuracy of fetal ultrasound for the prediction of absolute birth weight. No study has yet assessed if there is any direct relationship between fetal ultrasound macrosomia estimation and adverse outcome.

Although useful as a post birth definition, any strategy employed to identify and detect macrocosmic fetuses prior to birth will be limited by the imprecision of our measures of fetal weight. Until better measures of prediction are discovered and proven effective, we are dependent on these current ultrasound methods of fetal weight estimation. Therefore, the most

clinically useful definition of fetal macrosomia may in fact be based on antenatal prediction measurements and any possible association or predictive ability that they have with particular adverse outcomes. The question of which ultrasound measurement and estimated fetal macrosomia definition is most clinically useful for predicting adverse outcome is not answered by the currently available literature. This is the research question I seek to answer.

The antenatal prediction of fetal macrosomia is particularly pertinent now for two reasons. The 2015 Supreme Court ruling, Montgomery v Lanarkshire, has implications for the entire medical profession and has particularly highlighted the importance of counselling women with suspected fetal macrosomia. A recent Cochrane review (four trials, 1190 women) found that induction of labour at 37-40 weeks for suspected fetal macrosomia reduced birthweight (mean diff. 178g, 95% CI 41 to 315), fractures (RR 0.20, 95% CI 0.05 to 0.79) and shoulder dystocia (RR 0.60, 95% CI 0.37 to 0.98) (2). Given that we have an effective intervention for suspected fetal macrosomia, trying to improve the accuracy of our prediction of fetal macrosomia has never been more important.

3 Research planning and methods selection

3.1 Possible methods to approach this research

Although RCTs of a test strategy provide the most accurate, valid proof of association and prediction, it must first be known what tests are thought to be the most useful, to allow the hypothesis to be formulated and the RCT to be planned accordingly. Currently no published studies have systematically reviewed my selected index tests in relation to delivery and neonatal outcomes. A Cochrane review published in 2016 entitled "Regimens of fetal surveillance of suspected large-for-gestational-age fetuses for improving health outcomes" attempted to answer this question. They focused on the regimens of fetal surveillance including fetal movement counting, fetal heart rate assessment (cardiotocography), and the use of ultrasound for fetal growth scans, umbilical artery Doppler and assessing amniotic fluid volume. They found no RCTs that assessed the effect of antenatal fetal surveillance regimens of a suspected LGA fetus on important health outcomes for the mother and baby (151).

It is not known which ultrasound measurements provide the best guide to the prediction of adverse outcome. To enable future RCTs on this subject it is therefore of paramount importance that a systematic review of all current data is performed. Although there are no RCTs available to

answer my research question (151), there is however, a large amount of observational data in the form of cohort and case-control studies that have measured the different ultrasound prediction methods of fetal macrosomia and also recorded the outcomes of each mother and baby individually. It is by systematically reviewing this data and comparing the different ultrasound methods, that we seek to advise clinical practice and add to the knowledge base that allows future diagnostic and therapeutic RCTs to be planned. My data has the potential to provide a better-proven rationale for the selection of the index tests chosen to identify the participants who could benefit most from the interventions and should therefore be the patients enrolled in any future RCTs to assess interventions for predicted fetal macrosomia.

A comprehensive systematic review of the literature surrounding fetal macrosomia prediction by ultrasound will reveal and clarify the existence and strength of any relationship that may exist between the prediction tests and adverse delivery and neonatal outcomes. The results of the review will guide clinical practice regarding the choice of these different ultrasound tests and their ability to predict these adverse outcomes. This in turn will enable improved counselling and selection of the pregnancies that would benefit from preventative intervention. These studies are already in the recruitment stages due to the NIHR funding already offered for this cause (previously referred to in section 2.6.3).

3.2 Systematic reviews of diagnostic accuracy

High quality meta-analyses, systematic reviews of RCTs and RCTs with a very low risk of bias make up the highest level of evidence when considering evidence grades (152). We know that there is an absence of RCT data in the prediction of fetal macrosomia to outcome, and therefore systematic review of observational data is the best remaining possibility as a method to answer my research question. A systematic review aims to identify, appraise and interpret all the empirical evidence that meets pre-defined criteria to answer a specific research question (153). The methods required are systematic and explicit to produce reliable results. A diagnostic test accuracy review assesses how well a diagnostic test performs in diagnosing and detecting a particular disease or outcome. This is the method I have selected to best ask my research question and I will use the validated Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy to guide my methods (154).

I considered at the outset of my research planning what type of systematic review to pursue. Systematic reviews and meta-analyses that use individual patient data (IPD) have been described as the gold standard of systematic review (155). These reviews involve the central collection and new analysis of the original raw data from all studies worldwide that have addressed a common research question (156). However these studies take longer and are more resource intensive than other types of systematic review (156). I considered with my supervisors the

methodological factors that were likely to influence the results, in conjunction with the resource and time constraints. In addition, the fact that some of the literature for my review is quite old makes it less likely that it would be available for IPD-meta-analysis. The active decision made was that I would aim to collect data from published reports and collect any additional or replacement summary data from trialists but not to pursue an IPD meta-analysis.

3.3 Research questions addressed by this thesis

The outcome of this systematic review aims to contribute to the literature and advise clinical practice. It will be an original contribution as this is the first study to my knowledge to directly examine which antenatal imaging measures of fetal macrosomia best predict adverse maternal and neonatal outcomes. For clinicians needing to counsel women with suspected fetal macrosomia this is useful information.

The practice of obstetrics is outcome driven so this information will be highly clinically relevant. Ultimately I aim to answer: which current ultrasound measurement of predicted fetal macrosomia best correlates with adverse outcome and therefore indicates which women would potentially benefit most from early induction. This complements and potentially aids the further research recommendations of the Cochrane authors (2).

3.3.1 Research questions addressed in this thesis by systematic review

1. What is the association of USS predicted fetal macrosomia and adverse birth outcomes?
2. What is the predictive ability of USS estimated fetal macrosomia and adverse birth outcomes?
3. Does any association or predictive ability of estimated fetal macrosomia and adverse birth outcomes, differ within the diabetic and non-diabetic pregnant population?
4. Does any association or predictive ability of estimated fetal macrosomia differ between fetal and maternal adverse outcomes?

4 Methodology

4.1 Method for systematic review of which diagnostic definition of predicted fetal macrosomia best associates with adverse delivery and neonatal outcomes?

The systematic review performed in this thesis evaluates the association and any possible predictive ability of different measurements of antenatally predicted fetal macrosomia for adverse birth outcomes. This systematic review was conducted using a prospective protocol and common methodology from widely recommended sources (155, 157-160). This methodology includes the following stages:

1. Framing the review question and defining inclusion criteria
2. Study identification
3. Quality assessment
4. Data extraction
5. Statistical analysis and interpretation of data

4.1.1 Framing the question

Systematic reviews should always set a clear question at the outset. This helps to ensure that all objectives of the review are successfully

achieved. A review protocol is written out in advance, setting out the methods to be used with the aim of minimising bias. The review question is specified and the study inclusion and exclusion criteria are defined using the recommended 'PICOS' criteria (157). These criteria comprise the Population, Index test, Comparator, Outcome and Study design. In this study, the comparator criteria were not required or used. Using PICOS at the outset of the review allows the development of a well-structured question. This aids determination of the search strategy details and henceforth identifies the relevant studies for inclusion in the review.

The components of the question used in this review are described below.

1. **Question** Which measurement definition of predicted fetal macrosomia most accurately associates with and predicts adverse delivery and neonatal outcomes?

2. **Population** Inclusion criteria: Pregnant women who have had sonographic fetal measurements in the third trimester (at least 28 weeks gestation), that have been recorded as predicted fetal macrosomia **AND** that have outcome measures of their term delivery (≥37 completed weeks) also recorded. Exclusions include twin pregnancy or preterm delivery.

3. **Index test** Any US or MRI measurement defined as predicted fetal macrosomia including: EFW (>4kg, 4-4.5kg, >4.5kg, >90th centile, >95th centile); AC (>35cm, >36cm, >75th centile, >90th centile, >

95th centile); ratios of measurements and novel distances (e.g. AD/BPD, FL/AC, cheek to cheek, humeral soft tissue thickness).

4. **Outcome** Any reference standard looking at adverse delivery or neonatal outcome including shoulder dystocia, caesarean section rate, BPI, clavicle fracture, birth trauma, neonatal hypoglycaemia, postpartum haemorrhage, third or fourth degree perineal trauma.

5. **Study design** RCTs, test accuracy or epidemiological studies that consist of observational studies of defined populations. They must allow the generation of 2 x 2 tables (true positives, false positives, false negatives and true negatives) to calculate an estimate of the association and predictive ability between the index tests and reference standard outcomes. Cohort and case control studies will be used but case series of ≤10 will be excluded due to unreliability.

4.1.1.1 Population

The population for inclusion were pregnant women with singleton pregnancies that resulted in a term birth (≥37 weeks). These pregnant women must have had an index test of interest measured and recorded as predicted fetal macrosomia in the third trimester of pregnancy (≥28 weeks). Multiple pregnancies and premature deliveries were excluded due to the confounding effect of these cases on the adverse outcomes applied. No other restrictions (such as diabetes or other risk factors for adverse outcome) were applied but where possible these will be considered as subgroup analyses.

4.1.1.2 Index tests

All measurements and thresholds that have been described as methods to predict fetal macrosomia were included as index tests for the initial scoping searches (discussed in section 2.5). This diverse inclusion allowed the potential to capture the widest range of prediction data for all the modalities including two-dimensional ultrasound, three-dimensional ultrasound and MRI. Scoping searches were performed to determine whether there was available primary data for these measurement prediction parameters combined with the outcomes of interest. These initial searches revealed that for three-dimensional US and MRI studies there was only test accuracy data available and not outcome data recorded. Therefore, the index tests included in the review consisted of two-dimensional ultrasound data only. Although a comparison of the different modalities would be beneficial for development of future practice, as two-dimensional ultrasound is the modality that is currently widely available and used worldwide, the review outcome of two-dimensional ultrasound alone will be more widely applicable and immediately useful for current clinical practice.

4.1.1.3 Outcome measures

The many complications that are known to be associated with fetal macrosomia have been explored and discussed in section 2.3.2. The search was designed to optimise the identification and inclusion of all these possible adverse outcomes. By using open terms combined with specific adverse outcomes, the aim was to encompass and identify all studies that had recorded any of the associated fetal, maternal and early neonatal

adverse outcomes. The main outcome measures for this project were designated as BPI, shoulder dystocia and caesarean section. BPI was chosen as the fetal complication most likely to have long-term consequences. Shoulder dystocia and caesarean were chosen as common maternal complications that influence counselling and delivery planning decision making in clinical practice

4.1.1.4 Study design

A systematic review and meta-analysis of RCTs or observational studies (including prospective or retrospective cohort or case control designs) that allowed the generation of a 2x2 table to include true positives, false positives, false negatives and true negatives. The completion of this 2x2 table allowed calculation of an estimate of the association and predictive ability of the index test to the adverse outcomes, to compute an estimate of the association and predictive ability between test result and outcomes. Any studies with ≤10 cases were excluded due to the unreliability of a small sample. Case control studies, when participants are selected for inclusion on the basis of the presence of their adverse outcome, are known to be prone to spectrum bias (161). This fact was accounted for in the scoring system used in the quality assessment.

4.1.2 Study identification

I designed the initial scoping search to capture original studies that included data regarding the association between predicted fetal macrosomia and any adverse outcomes for the mother or infant at delivery

or in the early neonatal period. This was created using terms and elements from the question identified in section 3.1.1.

The search consisted of a combination of medical subject headings (MeSHs; e.g. fetal macrosomia, ultrasonography, prenatal, infant, caesarean section), keywords (e.g. birthweight, abdominal circumference, shoulder dystocia), word variants and alternative terms using the Boolean operator 'OR' to capture citations of relevant text. These were then combined using 'AND' with a combination of the MeSHs, keywords and word variants to capture the relevant outcomes. The search was restricted to human only studies but there were no language or study design restrictions applied. The MEDLINE search strategy was written first and then adapted for use in the other electronic databases (Appendix 2. Section 9.2). The search was planned to be performed by investigator R.C.R. (the thesis author). Databases were searched from their inception until the search date and the following sources were used to find the relevant literature.

1. MEDLINE (1966-July 2016)
2. EMBASE (1980-July 2016)
3. PubMed (no date restriction-July 2016)
4. The Cochrane Library
5. Sources of 'grey' literature: OpenGrey, Web of Science and contact with experts in the field
6. Hand searching of the specialist journals

7. Review article checking for references and included studies
8. Direct contact with authors of the manuscripts that looked promising for inclusion but I was unable to complete the 2x2 table from the published article or poster alone

The use of all of these sources combined, enabled coverage of the general biographical databases, specialist electronic databases, unpublished or 'grey' literature and conference proceedings.

A comprehensive database of collated citations was created using EndNote X7. The entire list of citations was scrutinised by the author and then the third reviewer (Dr N. Jones) repeated this selection process in either some or all of the citations depending on the number of citations found. Full articles were obtained for all citations that were felt, on examination of the title and abstract, to fulfil the predefined selection criteria. To prevent the exclusion of any relevant articles, papers were selected for full text further analysis if the title and abstract suggested any possibility of containing any two of the three essential criteria of Population, Index Test and Outcome measure. Translations of the articles that were published in languages other than English were obtained (see acknowledgements). The author (RCR) and the second reviewer (VW) performed a dual review of all full text papers. This was conducted by using a pre-defined data collection tool that enabled assessment of adherence to a checklist that included the strict PICOS criteria and the ability to populate 2x2 tables. All disagreements between the two reviewers had re-review

and if consensus could not be reached, a third reviewer, Dr N Jones' input was enlisted.

All manuscripts were carefully scrutinised to ensure no duplicate populations were used. When duplicates were identified, only the most complete and recent data sets were used for data extraction.

4.1.3 Quality Assessment

The methodological quality of all studies that were selected for inclusion in the systematic review were assessed. Two reviewers (RCR and VW) completed the same structured assessment independently. The STARD and QUADAS published checklists were used to create the quality assessment framework. These are both validated methods for the study of reporting and methodological quality, respectively, in the reporting of diagnostic test accuracy studies. (162, 163). From within these validated tools I selected the elements that best represented the methodical quality for systematic reviews that assess the association between prognostic tests and outcome. I used the same approach that has been used successfully by other similar, published systematic reviews and meta-analysis (164, 165). The study was rated by the following criteria (table 1):

Table 1: Criteria for quality assurance rating of studies included in the systematic review and meta-analysis of ultrasound prediction of suspected fetal macrosomia and association with adverse outcome

Cohort study design
Population/eligibility criteria adequately described
Prospective recruitment
Consecutive recruitment
Index test and outcome measure described in enough detail to allow replication
Assessors of the outcome measure blinded to the index test (were the reference standard results interpreted without knowledge of the results of the index test)
>90% of participants had the outcome measure recorded
A statement regarding the use of intervention between the index test and outcome
An appropriate interval between the index test and outcome

A study was rated as high quality if it met at least four criteria, as medium quality if three criteria points were reached and classed as low quality if two or less were observed.

Cohort study design is considered as superior to case control study design. The description of the population/eligibility criteria was considered adequate if the study had reported the gestational age at delivery as ≥37 weeks and the index test had been performed in the third trimester (≥28 weeks gestation). The setting from which the population was recruited was required to be recorded and any co-morbidities, risk factors or exclusions (multiple pregnancy) to be given. The population was considered unclear if the majority of this information was included but any single part of this information was missing. Prospective recruitment was deemed to create

less bias than retrospective study methods and was therefore considered superior. Consecutive or random recruitment was considered as the ideal. The index test (US measurement) was considered as described in enough detail if the method, the timing and the personnel who performed the test were all documented clearly. The outcome measure description was classed as adequate if the definition used for the condition in question was specified exactly, there was an adequate description of the threshold to meet this definition and there were details of who performed the assessment or identification to determine whether the mother or baby had the outcome of interest. The blinding of that individual to the index test result and any description of treatment that the individual had received in the interim, that may have affected the outcome, was also considered important. Verification bias was thought to be minimised by setting a completion of follow up of >90% of participants. This was assessed by calculating the number of patients receiving the outcome measure divided by the number receiving the index test, multiplied by 100.

The overall quality score of a study was not used to decide inclusion or exclusion from the review but where possible the meta-analysis or subgroup analysis would be performed according to study quality. The overall quality of the included studies and the adherence to each individual quality criteria will be presented in table form within the review.

4.1.4 Data extraction

Data was extracted using a standardised form, from the included studies, for assessment of study quality and evidence synthesis. The extracted information will include the methodology and timing of the index tests and reference standards that allow the generation of 2 x 2 tables (true positives, false positives, false negatives and true negatives) to calculate an estimate of the association between the index tests and reference standard outcomes. Two members of the review team (RCR and VW) performed the study selection and data extraction in duplicate.

The included study authors were contacted directly to request data if the checklist was completed successfully to match the PICOS pre specified criteria but the manuscript did not provide enough information to populate the 2x2 tables.

4.1.5 Statistical analysis and Interpretation of data

I used aggregate data to quantitatively synthesise odds ratios (ORs) and 95% confidence intervals (95% CIs) for each index test-outcome pair. The results were pooled for each index test using a random effect meta-analysis to reveal the potential prognostic association with outcome. If a strong and statistically significant prognostic association was identified between a test and an outcome measure, I planned to calculate sensitivity, specificity and likelihood ratios from the 2 x 2 tables and then use a bivariate random-effects meta-analysis model to assess the predictive ability of a test.

The prevalence was pooled for each outcome using a random-effects meta-analysis model and the exact binomial method. Sensitivity and specificity were pooled for each index test using a bivariate random-effects meta-analysis model. If the bivariate model could not be estimated or there were estimation concerns (e.g. for the correlation parameter), then univariate models were used to pool sensitivity and specificity separately, as recommended by Takwoingi et al. (159). The positive predictive value (PPV) for a new population was derived using the formula of Riley et al. (160) which uses the summary sensitivity, specificity and prevalence.

Any heterogeneity between the studies was dealt with in three ways. Firstly, by looking at how much the papers varied in study design. Statistically, by using the chi-square statistical test and visually by creating forest plots (including the estimated predictive interval) and hierarchical summary receiver-operating characteristic (HSROC) curves to explore the heterogeneity.

All analyses were performed in Stata (version 15).

Details of the protocol for this systematic review were prospectively registered on PROSPERO and can be accessed at www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42011001752 (Appendix 1. Section 9.1).

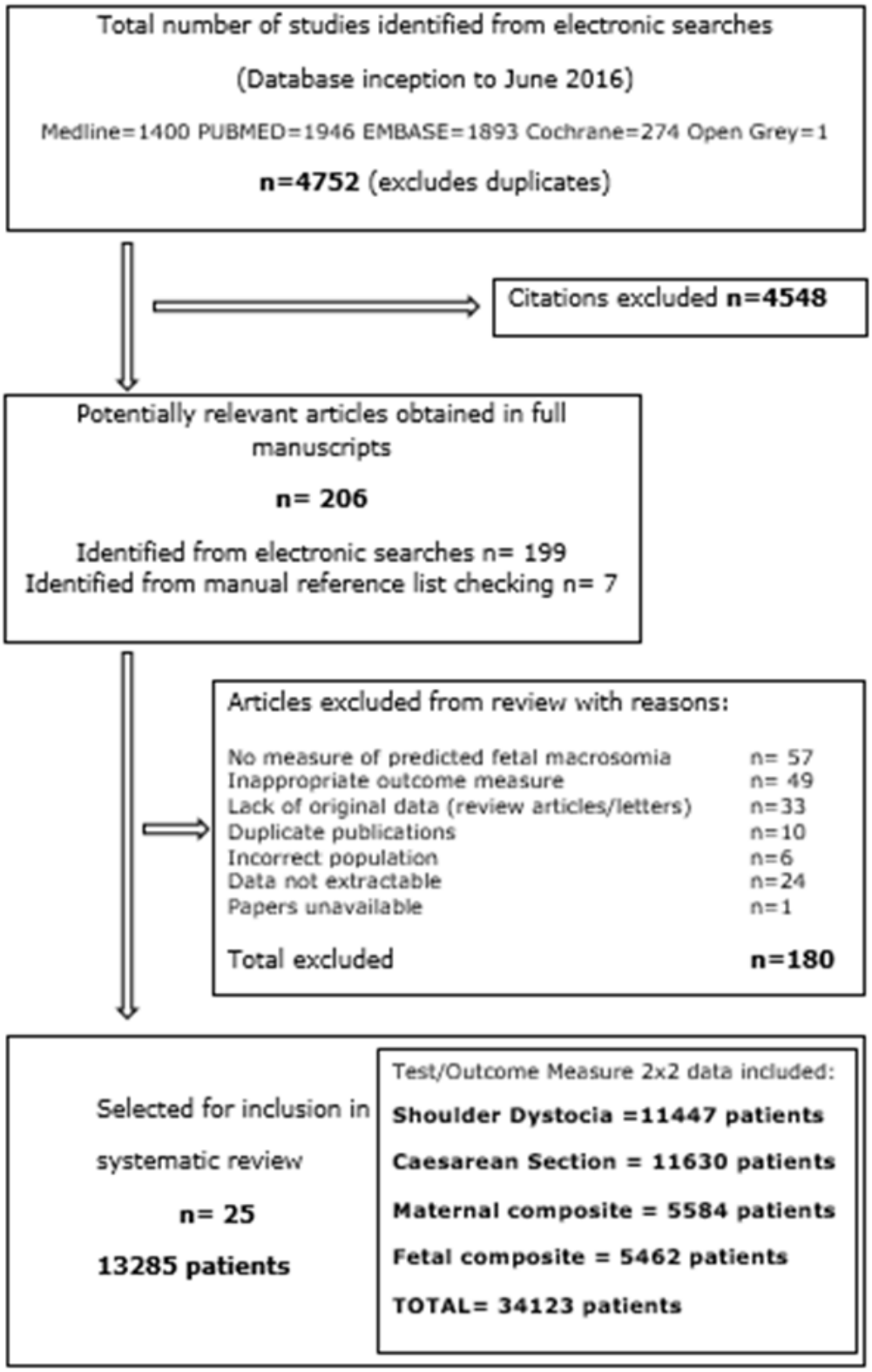
5 Results

5.1 **Study selection**

The total number of studies identified from the electronic searches was 4752. There were 206 articles obtained as full manuscripts, 199 of them identified from electronic searches and seven from manual reference list checking. The two reviewers used the pre-defined data collection tool to identify independently, articles that met the PICOS criteria and could populate 2x2 tables. This process excluded 180 of the full manuscript articles (see Appendix 4 for details and reasons for exclusion). Twenty-five observational studies (13,285 participants) were selected for inclusion (see Figure 1 below. Section 5.1.1). The characteristics of the included studies are shown in Table 5, Appendix 3. Twenty-two of the published studies included enough data within the publication to meet our inclusion criteria. A further three of these studies were included after contact with the authors, who provided further data to meet our inclusion criteria. A list of the excluded studies and the reasons for exclusion can be found in Table 6, Appendix 4. In total, after scrutiny of the published articles, there were another 23 articles that were considered to potentially contain the relevant data but not have presented it in a way that was accessible for me to extract for inclusion. These 23 authors were also approached but after contacting the authors they either did not respond, could not supply adequate data for creation of the 2x2 tables or clarification of the population resulted in exclusion.

5.1.1 Study selection process

Figure 1: PRISMA flow diagram for the systematic review and meta-analysis of ultrasound prediction of suspected fetal macrosomia and association with adverse outcome



5.2 Quality assessment

A study was rated as high quality if it met at least four criteria, as medium quality if three criteria points were reached and classed as low quality if two or less were observed.

According to our pre-specified criteria, 24 studies were classed high quality (96%) by meeting at least four criteria and one study by Basit (166) was classed as moderate quality (4%) meeting three criteria. The single paper that ranked as moderate quality was a retrospective cohort of 104 singleton pregnancies with newborns weighing 5kg or more. The definition of predicted macrosomia being tested was EFW>95% and the adverse outcomes recorded were shoulder dystocia, caesarean section rate and PPH.

The majority of the studies included were of cohort design (84%) and were retrospective (80%). There were, however, four prospective cohort studies and one prospective case control study. The details of the quality assessment of each individual study is included and presented within the following results chapters and in full below. Due to reporting methods there were lots of unclear assessment parameters.

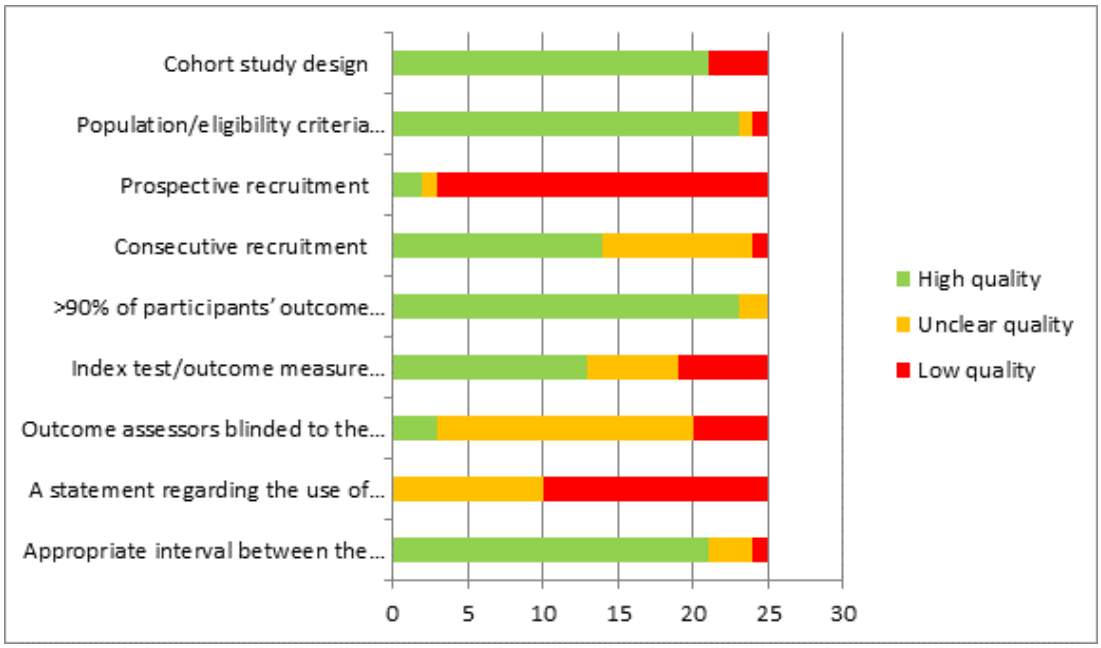
5.2.1 Quality assessment summary

The quality of all the included articles are summarised in Figure 2 (overleaf). Figure 3 summarises the quality of the studies grouped by criterion.

Figure 2: Quality assessment summary: review authors' judgements about each quality assurance item for each included study

	Cohort study design	Population/eligibility criteria adequately described	Prospective recruitment	Consecutive recruitment	Index test/outcome measure described in enough detail to allow replication	Outcome assessors blinded to the index test	>90% of participants' outcome measure recorded	A statement regarding the use of intervention between the index test and outcome	Appropriate interval between the index test and outcome
Aviram 2014	YES	YES	NO	?	?	NO	YES	NO	YES
Bailis 2010	YES	YES	NO	NO	?	?	YES	?	YES
Basit 2010	YES	YES	NO	?	NO	?	YES	NO	?
Bochner 1987	YES	YES	NO	?	YES	?	?	?	YES
Buikema 2014	YES	YES	NO	YES	NO	?	YES	NO	NO
Chaabane 2013	NO	YES	NO	YES	NO	?	YES	?	YES
Chauhan 2011	YES	YES	NO	YES	NO	?	YES	NO	YES
Chavan 2013	YES	NO	NO	YES	YES	NO	YES	NO	YES
Cohen 1996	YES	YES	NO	YES	YES	?	YES	?	YES
Cohen 1999	YES	YES	NO	YES	YES	YES	YES	?	YES
Eken 2015	NO	YES	?	YES	YES	?	YES	NO	?
Elliott 1982	YES	YES	NO	?	?	?	YES	NO	YES
Endres 2015	NO	YES	NO	YES	YES	?	YES	NO	YES
Kehl 2011	YES	YES	YES	YES	YES	?	YES	NO	YES
Levine 1992	YES	YES	NO	?	YES	?	YES	?	YES
Melamed 2010	NO	YES	NO	YES	?	NO	YES	NO	YES
Miller 2007	YES	YES	NO	YES	YES	YES	YES	NO	YES
Nelson 2009	YES	YES	NO	?	YES	?	YES	NO	YES
Parry 2000	YES	YES	NO	?	YES	NO	YES	NO	YES
Peleg 2015	YES	YES	NO	YES	?	NO	YES	?	YES
Rajan 2009	YES	YES	NO	?	YES	YES	YES	?	YES
Scifres 2015	YES	YES	NO	YES	YES	?	YES	NO	YES
Simpson 2015	YES	YES	NO	?	NO	?	YES	?	YES
Tatarova 2004	YES	?	NO	?	?	?	?	?	?
Weiner 2002	YES	YES	YES	YES	NO	?	YES	NO	YES

Figure 3: Quality assessment summary: review authors' judgements for all included studies, grouped by criteria



5.3 Whole population

5.3.1 Geographical location

The studies included in the systematic review encompassed populations from seven different countries worldwide. Almost two thirds of the studies were conducted in the USA. The distribution of the 25 included studies and populations, by country, were 16 USA (7270 patients), four Israel (6015 patients), one Ireland (104 patients), one Tunisia (930 patients), one Turkey (28 patients), one Germany (258 patients) and one Bulgaria (1847 patients). They were all written in English except for the Bulgarian article that required translation.

5.3.2 Diabetes status

The systematic review population included women with diabetes and without diabetes. Many (76%) of the included studies documented the diabetes status of their individual study populations. Three of the study populations exclusively had diabetes (including pre-existing and GDM), four studies had GDM only, seven studies had a mixed (background) rate of diabetes, four studies had no women with diabetes, one study matched the diabetes and non-diabetes populations and six studies did not record the diabetes status.

5.3.3 Gestation at scan and timing in relation to delivery

All ultrasound scans were performed in the third trimester. There was some variation between the studies of the gestation at scan and the timing in relation to delivery. Two of the studies purely specified the timing as being third trimester without any further detail. Four studies performed the predicted macrosomia ultrasound between 28-32w and one study between 28-36w. The other 18 studies all performed the predicted macrosomia ultrasound scan between 36w gestation and the subsequent term delivery. Eleven of these 18 studies additionally defined the time period that elapsed between the scan and the birth. These defined time periods between the ultrasound scan and delivery were documented as one study during labour, four studies at three days, one study at one week, three studies at two weeks and two studies at four weeks prior to delivery.

5.3.4 Similarities and differences between the populations

All the women included in the study were from tertiary care centres. The data was collected from a single centre population in 24 of the studies, whilst one study was multicentre from six centres.

The studies included were from a range of times in history. The date range of the inclusions was 1982-2015. There were two studies from 1980-89, three studies from 1990-1999, six studies from 2000-2009 and 14 (54%) from 2010-2015.

5.4 Main outcomes: Brachial Plexus Injury, Shoulder dystocia, Caesarean section

5.4.1 Brachial Plexus Injury

Three studies (131, 167, 168) were found to be suitable for inclusion that recorded BPI as an outcome. Each of these three studies used a different test measure; therefore, no meta-analysis was possible. There was a total of 909 patients included in these studies.

The first study, authored by Endres (131) was a multicentre, matched, retrospective, case-control study. The population was 362, as matched pairs of 181 participants who either did or did not experience vaginal delivery complicated by shoulder dystocia (also matched for the number of patients with diabetes). The ultrasound was performed within 4 weeks of a delivery at a gestation of greater than 36 weeks. The test measure was an AC to HC difference of greater than or equal to 50mm. This study

was assessed to be of high quality due to adequately described population/eligibility criteria, consecutive recruitment, index test and outcome measure described in enough detail to allow replication, >90% of participants had the outcome recorded, and an appropriate interval between the index test and outcome. There was a significant association found between an AC to HC difference of greater than or equal to 50mm and BPI: Odds ratio 7.2 (95% CI 1.78-29.0). This was not found to be a sensitive test measure: Sensitivity 0.23 (95% CI 0.00-0.46) but the specificity was very high: Specificity 0.96 (95% CI 0.94-0.98). These findings resulted in a high positive likelihood ratio of 5.75 and a lesser negative likelihood ratio of 0.80.

The second study, authored by Cohen (167), was a small, retrospective cohort study of 31 participants. All of the included population had diabetes. The test measure was performed within two weeks of delivery and the measure was abdominal diameter minus biparietal diameter, greater than or equal to 2.6cm. The study was assessed to be of high quality due to cohort study design, adequately described population/eligibility criteria, consecutive recruitment, index test and outcome measure described in enough detail to allow replication, >90% of participants had the outcome recorded, and an appropriate interval between the index test and outcome. There was no association found between AD-BPD≥ 2.6cm and BPI: Odds ratio 1.8 (95% CI 0.06-47.1).

The third study, authored by Weiner (168) was a prospective cohort study with 315 participants. This population included a background rate of diabetes of 9.2%. The test measure ultrasound was performed within three days of delivery. The measure recorded was EFW greater than or equal to 4000g using the Shepard formula. The study was assessed to be of high quality due to cohort study design, adequately described population/eligibility criteria, prospective recruitment, consecutive recruitment, index test and outcome measure described in enough detail to allow replication, >90% of participants had the outcome recorded, and an appropriate interval between the index test and outcome. There was no association found between EFW≥4000g Shepard and BPI: Odds ratio 1.4 (95% CI 0.08-21.8).

A summary of the BPI individual study results are shown in table 2.

Table 2: Brachial plexus injuries. The odds ratio and 95% confidence intervals by ultrasound measurement

Study	**Study size**	**Predicted macrosomia definition**	**Odds ratio**	**95% Confidence intervals**
Endres, 2015	362	AC to HC difference of >=50mm	7.2	(1.8 to 29.0)
Cohen, 1996	31	AD-BPD>or=2.6cm	1.8	(0.07 to 47.1)
Weiner, 2002	315	EFW>=4000g Shepard	1.4	(0.08 to 21.8)

5.4.2 Shoulder dystocia

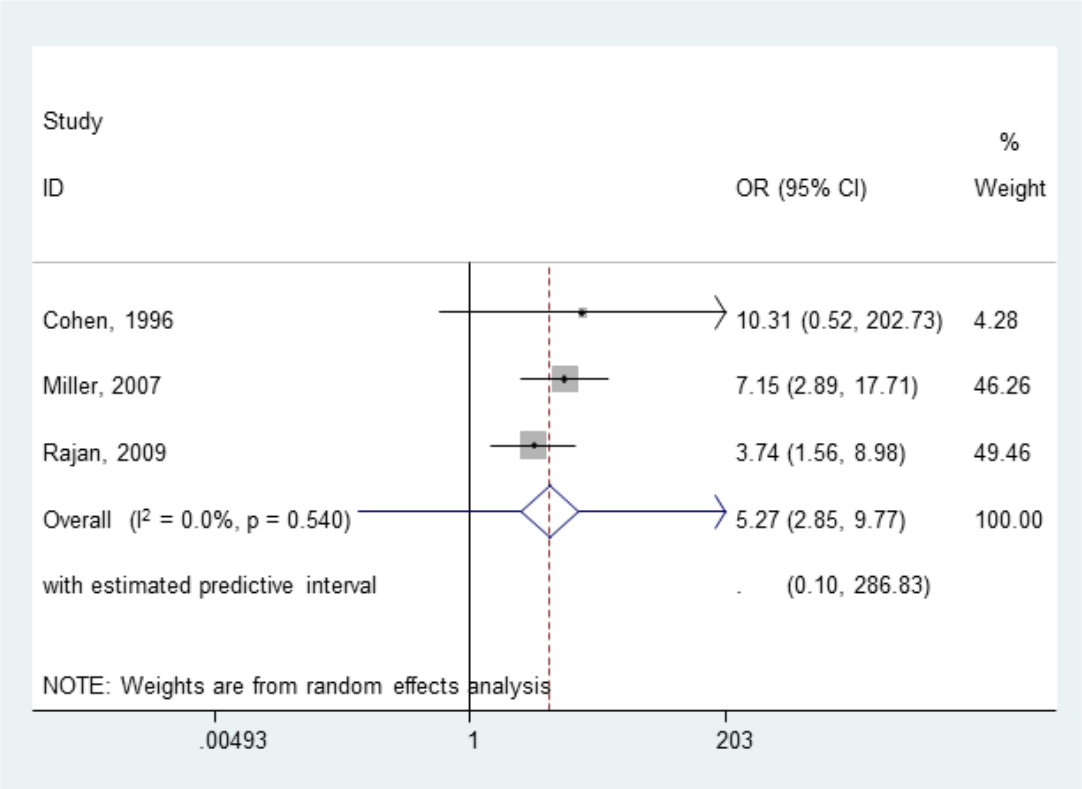
The outcome shoulder dystocia was identified in 18 of the included studies with a total of 11447 patients (112, 130-135, 166-176). From these 18 studies there were 26 'test measure to outcome' data pairs identified, as four of the studies recorded multiple test measures for their shoulder dystocia outcomes (Bailis(130); three, Eken(174); three, Miller(133); four, Cohen(167); two and Rajan(134); two).

Shoulder dystocia, meta-analysis was possible for three of the test measures. Meta-analysis is only possible when there are a minimum of three studies measuring and assessing the same test measure. Three different test measures met these conditions to allow meta-analysis to proceed, namely, EFW greater than 4000g, AC greater than the 90th centile and an AD-BPD greater than or equal to 2.6cm. These meta-analyses included four, three and three studies, respectively. The positive predictive values were predicted for a new population of these ten studies from a summary prevalence of 4.9% (95% CI 2.1-11.2) for shoulder dystocia.

The first meta-analysis was to investigate if there was any significant association demonstrated between an 'AD-BPD of greater than or equal to 2.6cm' and the outcome shoulder dystocia. The three studies included in this meta-analysis were Cohen, Miller and Rajan (133, 134, 167). The combined total of patients was 512. There was a significant association found between AD-BPD greater than or equal to 2.6cm and shoulder dystocia. This test resulted in the highest odds ratio of 5.27 (95% CI 2.85-

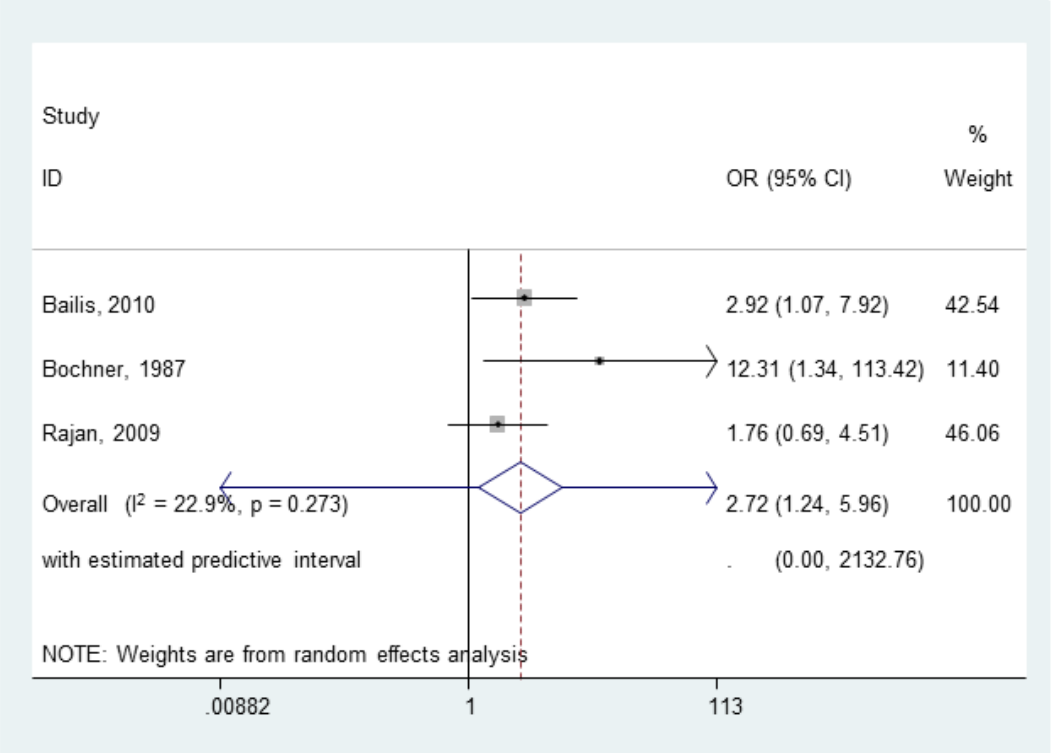
9.77) (Figure 4). The sensitivity of the test was 66.6 (95% CI 35.9=87.7) and the specificity was 71.5 (95% CI 44.0-88.9). This resulted in a positive likelihood ratio of 10.8.

Figure 4: Forest plot for odds ratio and 95% confidence intervals for shoulder dystocia detected by AD-BPD ≥2.6cm



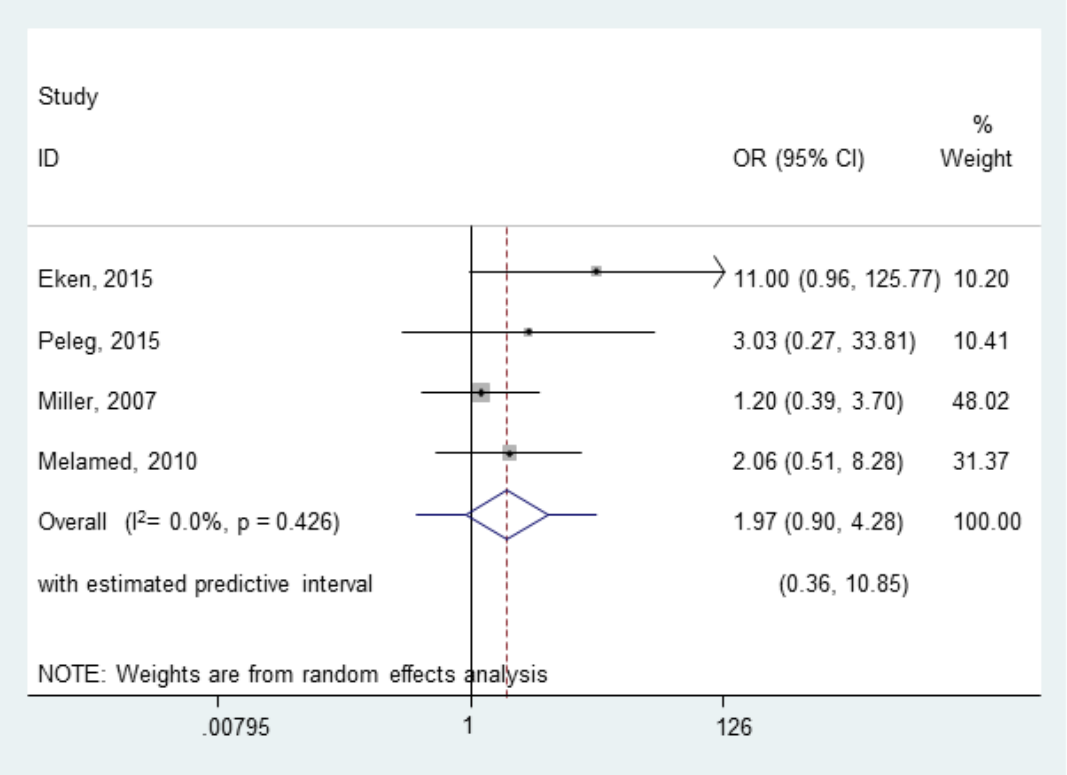
The second meta-analysis was to investigate if there was any significant association demonstrated between an abdominal circumference greater than the 90th centile and the outcome shoulder dystocia. There were three studies included in this meta-analysis, Bailis, Bochner and Rajan (130, 134, 170). The combined total of patients was 519. There was a significant association found between AC>90th centile and shoulder dystocia: Odds ratio 2.72 (95% CI 1.24-5.96) (Figure 5). The test had a sensitivity of 64.6 (95% CI 50.2-76.7) and a specificity of 64.7 (95% CI 45.3-80.3). The positive likelihood ratio of AC>90th centile was 8.6.

Figure 5: Forest plot for odds ratio and 95% confidence intervals for shoulder dystocia detected by AC >90th centile



The third meta-analysis was to investigate if there was any significant association demonstrated between an estimated fetal weight of greater than 4000g and the outcome shoulder dystocia. This combined the results of the 4 studies authored by Melamed (177), Eken (174), Peleg (178) and Miller (133) and resulted in 2731 patients for the meta-analysis. A non-significant association was found between EFW>4000g and shoulder dystocia with an odds ratio of 1.97 (95% CI 0.90-4.28) (Figure 6).

Figure 6: Forest plot for odds ratio and 95% confidence intervals for shoulder dystocia detected by EFW>4000g



5.4.3 Shoulder dystocia individual study results

Eight studies (130-134, 170-172) found a positive and significant odds ratio between the test measures and shoulder dystocia. The studies by Bochner and Cohen (1999) contained a population exclusively with gestational diabetes.

Bochner authored a retrospective cohort study of 201 patients. The test measure was an abdominal circumference of greater than the 90th centile on ultrasound performed at between 30-33w gestation. The study was assessed to be of high quality due to meeting four of the pre-defined quality assurance criteria (Fig 2. Section 5.2.1). There was a significant association found between AC>90th centile and shoulder dystocia: Odds

ratio 12.3 (95% CI 1.335-113.4). The test had high sensitivity 0.80 (95% CI0.45-1.15) and specificity 0.75 (95% CI 0.69-0.82) resulting in a positive likelihood ratio of 3.26 and a negative likelihood ratio of 0.27.

The 1999 study by Cohen was a retrospective cohort of 123 patients. The test measure was a residual AD-BPD of 0.33-2.0cm (actual AD-BPD minus 50th centile AD-BPD of the sample). The measure was performed at the most recent ultrasound after 30 weeks (at an average of 36 weeks gestation). This study gained the joint highest quality assessment of the systematic review due to meeting seven of the pre-defined quality assurance criteria (Fig 2. Section 5.2.1). A significant association was found between a residual AD - BPD of 0.33-2.0cm and shoulder dystocia with an odds ratio of 3.0 (95% CI 1.24-7.4). The sensitivity was 0.54 (95% CI 0.35-0.73) and the specificity was 0.72 (95% CI 0.63-0.81), resulting in a positive likelihood ratio of 1.93 and a negative likelihood ratio of 0.64.

The positive and significant study findings from Bailis (130), Endres (131) and Miller (133) were from populations that stated a "background rate" of diabetes in their included patients.

Bailis (130) authored a retrospective cohort study with 159 participants. There were three test measures assessed for the outcome of shoulder dystocia: AC>90th centile, HC<50th centile and a HC/AC<0.95. The test measures were performed at a gestation greater than 30 weeks. The study

was assessed to be of high quality due to meeting four of the pre-defined criteria (Fig 2. Section 5.2.1). There was a significant association found between AC>90th centile and shoulder dystocia with an odds ratio of 2.9 (95% CI 1.08-7.9). The sensitivity of the test was 0.5 (95% CI 0.27-0.73) and the specificity was 0.74 (95% CI 0.67-0.82). The resultant positive likelihood ratio was 1.96 and the negative likelihood ratio was 0.67. In the same study population there was no association found between HC< 50th centile and shoulder dystocia; Odds ratio 2.8 (95% CI 0.96-8.4) or between HC/AC<0.95 and shoulder dystocia; Odds ratio 1.4 (95% CI 0.52-3.8).

The study authored by Endres (131) (detailed in 5.4.1) had a test measure of AC-HC≥ 50mm. There was a significant association found between AC-HC≥50mm and shoulder dystocia with an odds ratio of 7.7 (95% CI 1.73-34.1). The sensitivity was very low at 0.08 (95% CI 0.04-0.12) but the specificity was very high at 0.99 (95% CI 0.97-1.00). This gave a high positive likelihood ratio of 7.13 but not a good negative likelihood ratio of 0.93.

The study authored by Miller was a retrospective cohort of 322 patients. They stated the background rate of diabetes as 19%. Ultrasound was performed within 2 weeks of delivery and there were four test measures assessed. AD-BPD≥2.6cm, AC>95th centile, EFW>97th centile and EFW>4000g. This study was the joint highest quality assessment of the systematic review due to meeting seven of the pre-defined criteria (Fig 2.

Section 5.2.1). There was a significant association found between AD-BPD≥ 2.6cm and shoulder dystocia: Odds ratio 7.2 (95% CI 2.89-17.7). The sensitivity was 0.43 (95% CI 0.23-0.64) and the specificity 0.90 (95% CI 0.87-0.94) resulting in a positive likelihood ratio of 4.48 and a negative likelihood ratio 0.63. The same study demonstrated no significant association between shoulder dystocia and AC>95th centile; odds ratio 2.0 (95% CI 0.81-4.9), EFW>97th centile; odds ratio 2.9 (95% CI 0.79-11.0) or EFW>4000g; odds ratio 1.2 (95% CI 0.39-3.7).

Rajan (134) authored the only study that found a significant positive association in a population with no incidence of gestational or pre-existing diabetes.

The Rajan (134) study was a retrospective cohort of 159 patients. Ultrasound was performed at greater than or equal to 36 weeks gestation and all patients had an EFW of greater than 4kg. There were two test measures assessed: AD-BPD≥2.6cm and AC>90th centile. This study was assessed to be of high quality due to meeting six of the pre-defined criteria (Fig 2. Section 5.2.1). There was a significant association found between AD-BPD≥ 2.6cm and shoulder dystocia with an odds ratio of 3.7 (95% CI 1.56-9.0) a sensitivity of 0.67 (95% CI 0.49-0.84) and a specificity of 0.65 (95% CI 0.57-0.73). The positive likelihood ratio was 1.91 and the negative likelihood ratio was 0.51. There was no association found between AC>90th centile and shoulder dystocia: Odds ratio 1.8 (95% CI 0.69-4.5).

Two (171, 172) of the studies that demonstrated a positive and significant association between the test measures and shoulder dystocia, did not know or record the diabetes status of their population.

The study authored by Buikema (171) was a retrospective cohort of 247 patients undergoing ultrasound at between 37-42w gestation. The test measure was EFW> 90th centile (the entire cohort had an abdominal circumference of greater than 35cm). This study was assessed to be of high quality due to meeting four of the pre-defined criteria (Fig 2. Section 5.2.1). There was a significant association found between EFW>90th centile and shoulder dystocia with and odds ratio of 6.2 (95% CI 2.29-16.7). The sensitivity was 0.38 (95% CI 0.17-0.59) and specificity was 0.91 (95% CI 0.87-0.95) with a positive likelihood ratio of 4.21 and negative likelihood ratio of 0.68.

Chaabane (172) authored a retrospective case control study with 930 patients. Ultrasound scan was performed between 37-42 weeks gestation, all within 72hrs of delivery. The index test measure was an AC≥ 350mm. This study was assessed to be of high quality as it met four of the pre-defined criteria (Fig 2. Section 5.2.1). There was a significant association found between AC ≥350mm and shoulder dystocia: Odds ratio 5.9 (95% CI 3.20-10.8). The sensitivity; 0.84 (95% CI 0.77-0.92) was found to be higher that the specificity; 0.52 (95% CI 0.49-0.56). This gave a positive likelihood ratio of 1.76 and a negative likelihood ratio of 0.30.

The ten (135, 166-169, 173, 174, 176-178) other studies did not demonstrate a significant association between their test measures and shoulder dystocia.

Four of these studies, by Scifres (176) and Chavan (173) had populations with GDM. Elliott (135) and Cohen, 1996 (167) had populations of patients with a mixture of pre-existing and GDM.

Scifres (176) authored a retrospective cohort study. There were 903 participants. Ultrasound was performed within 31 days of delivery and the test measure recorded was EFW>90th centile. This study was assessed to be of high quality as it met six of the pre-defined criteria (Fig 2. Section 5.2.1). There was no significant association found between AC>90th centile and shoulder dystocia: Odds ratio 1.6 (95% CI 0.61-4.0).

Chavan (173) was the author of a retrospective cohort study of 296 patients. Ultrasound was performed at 28-36w gestation. The index test measure was AC≥ 75th centile. This study was assessed to be of high quality due to meeting five of the pre-defined quality criteria (Fig 2. Section 5.2.1). There was no association found between AC≥75th centile and shoulder dystocia: Odds ratio 0.7 (95% CI 0.12-3.8).

Elliott (135) authored a retrospective cohort study of 70 patients. The ultrasound was performed within 2 weeks of delivery. The test measure assessed was a Chest diameter (CD)-BPD≥ 1.4cm. This study was assessed to be of high quality due to meeting four of the pre-defined criteria (Fig 2. Section 5.2.1). There was no significant association found between CD-BPD≥1.4cm and shoulder dystocia: Odds ratio 3.6 (95% CI 0.36-36.4).

The small, 31 patients, 1996 study (167) authored by Cohen (study detailed in 5.4.1) performed AD-BPD≥2.6cm within two weeks of delivery. There was no significant association found between AD-BPD≥ 2.6cm and shoulder dystocia. Although the odds ratio was found to be high: Odds ratio 10.3 (95% CI 0.52-202.7), the small study sample did not allow the positive association to reach significance.

Two studies (168, 174) that found no association used patients with a "background rate "of diabetes.

The study by Eken (174) was a small, prospective cohort study of 28 patients. All patients delivered babies, between the years of 2010-2014 in a tertiary referral centre that resulted in a brachial plexus injury. The ultrasound scan was performed in the late third trimester. There were three index test measures recorded: EFW> 4kg, AC> 350mm and an AC>360mm. This study was assessed to be of high quality as it met four

of the pre-existing quality criteria (Fig 2. Section 5.2.1). There was no significant association found between any of the three test measures. For EFW >4kg and shoulder dystocia, although the odds ratio of 11.0 (95% CI 0.96-125.8) was high, the small number of study participants made the confidence interval too wide for significance to be found. There was no significant association found for AC>350mm or for AC>360mm and shoulder dystocia. The odds ratios were 2.1 (95% CI 0.34-13.4) and 1.5 (95% CI 0.29-7.7) respectively.

The study authored by Weiner (168) was a prospective cohort study with 315 participants. This population stated the background rate of diabetes as 9.2%. The test measure ultrasound was performed within three days of delivery. The measure recorded was EFW≥ 4000g using the Shepard formula. The study was assessed to be of high quality due to meeting six of the pre-defined criteria (see Figure 2). There was no significant association found between EFW≥4000g and shoulder dystocia the odds ratio was 2.7 (95% CI 0.25-30.4).

Two studies (177, 178) excluded all patients with diabetes and did not demonstrate any significant results.

Melamed (177) authored a large retrospective case-control study of 1938 patients. All patients delivered a live born fetus weighing between 3500g -4499g. Ultrasound was performed in all within three days of the delivery.

The test measure was EFW= 4000-4999g. This study was assessed to be of high quality due to meeting four of the pre-defined criteria (Fig 2. Section 5.2.1). There was no significant association found between EFW=4000-4999g and shoulder dystocia; Odds ratio 2.1 (95% CI 0.51-8.3)

Peleg (178) authored a retrospective cohort study of 443 patients who delivered an infant of greater than 4000g in the year of 2012. All patients were ultrasound scanned at 38-42 weeks on admission, in labour. The ultrasound test measure was EFW≥4kg. This study met five of the quality assurance criteria (Fig 2. Section 5.2.1) so was assessed to be of high quality. No association was found between EFW≥ 4kg and shoulder dystocia. The odds ratio calculated was 3.0 (95% CI 0.27-33.8).

The final two studies that did not demonstrate a positive and significant association were Aviram (169)and Basit (166). Neither of these studies knew or recorded the diabetes status of the patients included.

The study authored by Aviram (169) was a large retrospective cohort study of 3319 participants. The scans were performed at greater than 37 weeks gestation and the test measure was HC>90th centile (>347mm). The study met four of the pre-defined criteria and was therefore assessed to be of high quality (Fig 2. Section 5.2.1). There was no association found between

a HC> 90th centile and shoulder dystocia: Odds ratio 3.3 (95% CI 0.90-12.4).

The study authored by Basit (166) was a retrospective cohort of patients with singleton pregnancies that resulted in the birth of macrosomic infants weighing greater than 5kg. There were 104 patients that had an antenatal ultrasound scan that predicted the EFW. The index test in this study was an EFW >95th centile. The study was assessed to be of medium quality as it only met three of the pre-specified criteria (Fig 2. Section 5.2.1). There was no association found between EFW>95th centile and shoulder dystocia: Odds ratio 0.9 (95% CI 0.084-10.2).

Table 3 summarises all of the shoulder dystocia systematic review results of the individual included articles.

Table 3: Shoulder dystocia: The odds ratio and 95% confidence intervals by ultrasound measurement

Study	Study size	Predicted macrosomia definition	Odds ratio	95% Confidence intervals
Eken, 2015	28	AC>350mm	2.1	0.34 to 13.4
Eken, 2015	28	AC>360mm	1.5	0.29 to 7.7
Bailis, 2010	159	AC>90th centile	2.9	1.08 to 7.9
Bochner, 1987	201	AC>90th centile	12.3	1.34 to 113.4
Rajan, 2009	159	AC>90th centile	1.8	0.69 to 4.5
Miller, 2007	322	AC>95th centile	2.0	0.81 to 4.9
Chaabane, 2013	930	AC>or=350mm	5.9	3.20 to 10.8
Chavan, 2013	296	AC>or=75th centile	0.7	0.12 to 3.8
Cohen, 1996	31	AD-BPD>or=2.6cm	10.3	0.52 to 202.7
Miller, 2007	322	AD-BPD>or=2.6cm	7.2	2.89 to 17.7
Rajan, 2009	159	AD-BPD>or=2.6cm	3.7	1.56 to 9.0
Elliott, 1982	70	Chest diameter-BPD>=1.4cm	3.6	0.36 to 36.4
Melamed, 2010	1938	EFW =4000-4999	2.1	0.51 to 8.3
Eken, 2015	28	EFW>=4000g Hadlock	11.0	0.96 to 125.8
Peleg, 2015	443	EFW>=4000g Hadlock	3.0	0.27 to 33.8
Weiner, 2002	315	EFW>=4000g Shepard	2.7	0.25 to 30.4
Miller, 2007	322	EFW>4000g	1.2	0.39 to 3.7
Buikema, 2014	247	EFW>90th centile (all AC.35cm)	6.2	2.29 to 16.7
Scifres, 2015	903	EFW>90th centile Hadlock	1.6	0.61 to 4.0
Basit, 2010	104	EFW>95th centile	0.9	0.08 to 10.2
Miller, 2007	322	EFW>97th centile	2.9	0.79 to 11.0
Bailis, 2010	158	HC/AC<0.95	1.4	0.52 to 3.8
Bailis, 2010	158	HC<50th centile	2.8	0.96 to 8.4
Aviram, 2014	3319	HC>90th centile (>347mm)	3.3	0.90 to 12.4
Cohen, 1999	123	Residual AD-BPD (actual-50th centile)=0.33-2.0cm	3.0	1.24 to 7.4

5.4.4 Caesarean Section

Fifteen studies reported caesarean section as an outcome. They included 11,630 women test/outcome data sets. The summary prevalence for caesarean section in the studies was 36% (95% CI 23-51%). There were 15 'test measure to outcome data pairs' identified although one study recorded multiple test measures for their caesarean section outcomes (179) and two studies (175, 180) recorded results for total caesareans and a group excluding elective caesareans.

Caesarean section meta-analysis was possible for two test measures; an EFW≥ 4000g and an EFW≥ 90th centile. These two meta-analyses included five and three studies respectively. Unfortunately, meta-analysis was not possible for any of the AC or compound measurements as there were not the minimum of three studies testing these measures, which are needed to complete meta-analysis. There were therefore no meta-analysis results to compare to these EFW test measure meta-analysis results.

The first meta-analysis was to investigate if there was any significant association demonstrated between an EFW≥ 4000g and caesarean section. The five studies included in this meta-analysis were Melamed(177), Chauhan (179), Tatarova (181), Parry (182) and Peleg (175). The total number of patients with these five studies combined was of 4569. There was a significant association found between EFW≥4000g and caesarean section with an odds ratio of 4.98 (95% CI 3.07-8.06) (Figure 7). The sensitivity was calculated as 52.1 (95% CI 28.7-74.6) and the specificity

was 82.3 (95% CI 60.3-93.5) (Figure 8). This gave a positive likelihood ratio of 62.2.

Figure 7: Forest plot for odds ratio and 95% confidence intervals for caesarean section detected by EFW ≥4000g

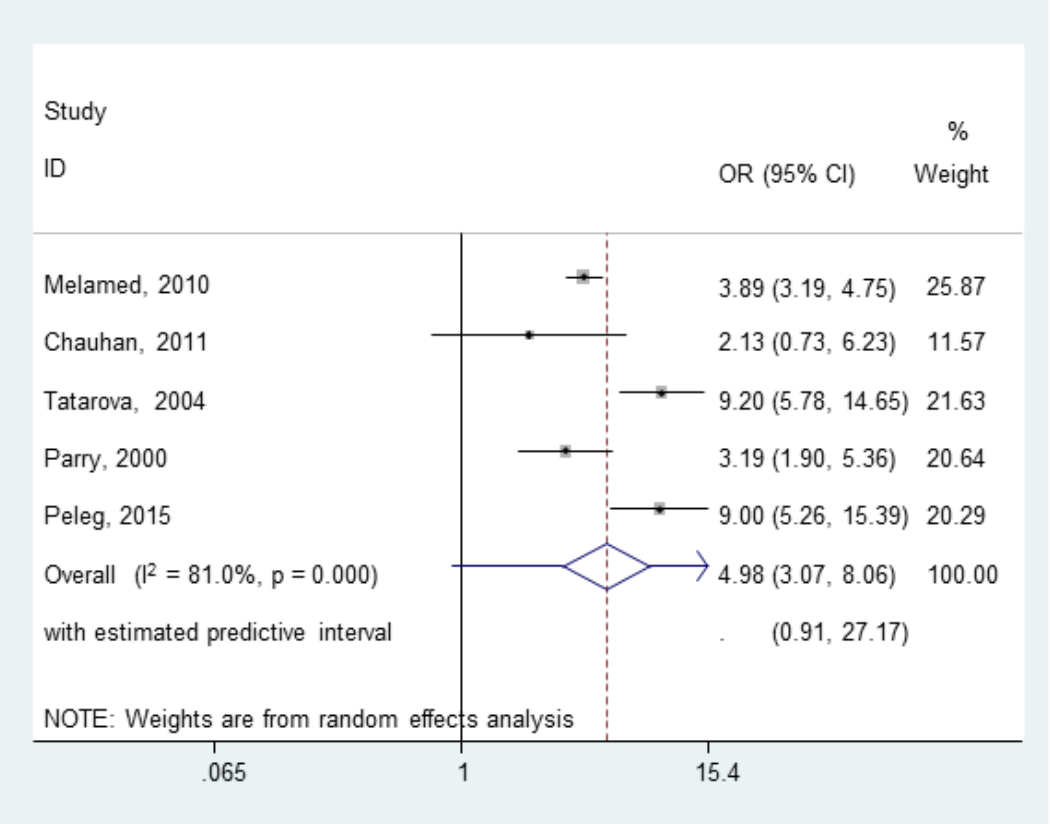
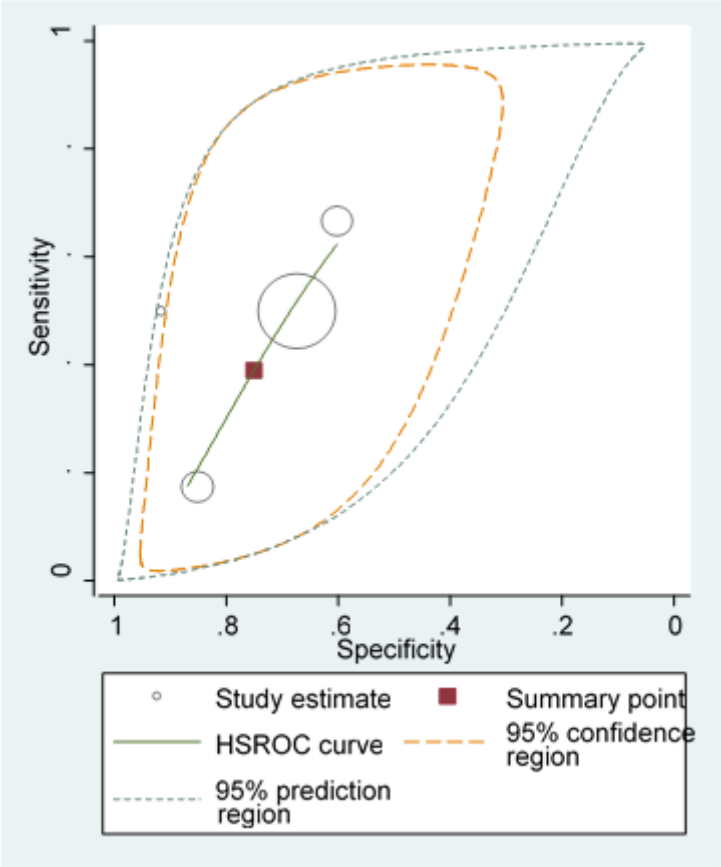


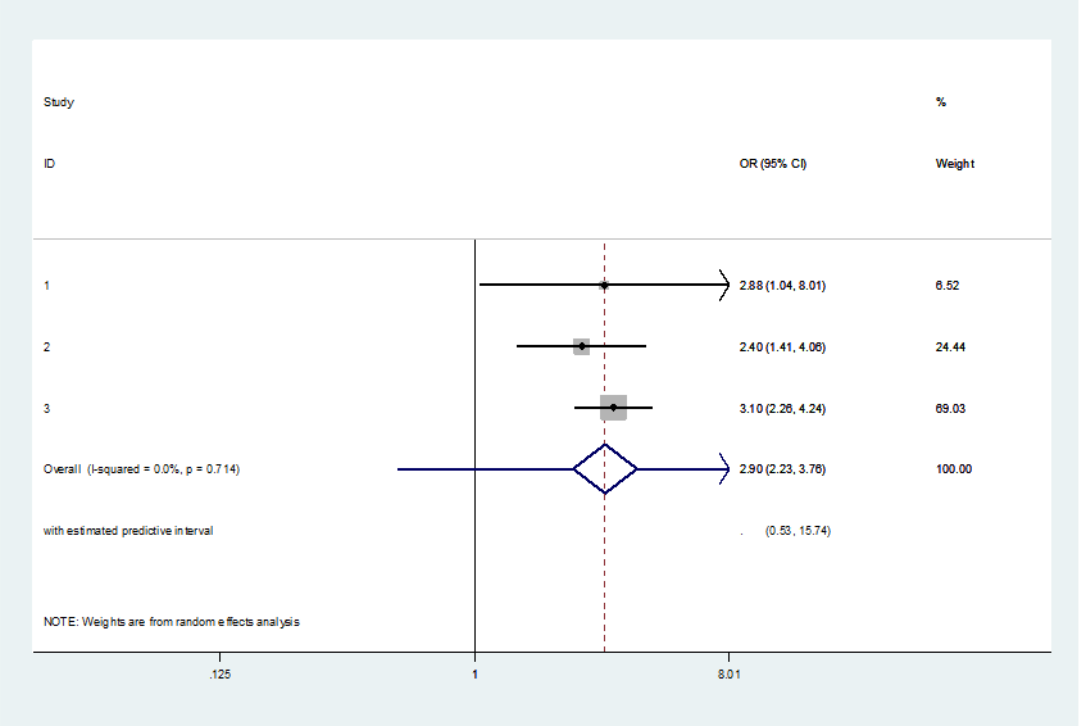
Figure 8: ROC plot for odds ratio and 95% confidence intervals for caesarean section detected by EFW ≥4000g



The second meta-analysis was to investigate if there was any significant association between an EFW≥ 90th centile and caesarean section. The three studies included in this meta-analysis were Chauhan (179), Levine (183) and Scifres (176). These three studies had a combined total of 1386 patients. There was a positive significant association found between EFW≥90th centile and caesarean section with an odds ratio of 2.90 (95% CI 2.23-3.76) (Figure 9).

Figure 9: Forest plot for odds ratio and 95% confidence intervals for caesarean section detected by EFW>90th centile

(Note: 1= Chauhan, 2011; 2= Levine, 1992; 3= Scifres, 2015)



5.4.5 Caesarean section individual study results

There were two studies, Chavan (173) and Scifres (176) who's population's consisted exclusively of patients with diabetes and that yielded positive and significant associations,. Both of these studies included only patients with GDM and not pre-existing diabetes.

Chavan (173) authored a retrospective cohort study of 296 patients (study detailed in 5.4.3). There was a significant association found between an AC≥75th centile and caesarean section. The odds ratio was 1.7 (95% CI 1.01-3.0). Sensitivity was calculated as 0.52 (95% CI 0.40-0.64) and

specificity as 0.61 (95% CI 0.55-0.68) giving a positive likelihood ratio of 1.35 and a negative likelihood ratio 0.78.

The study authored by Scifres (176) was a retrospective cohort study with 903 participants (study detailed in 5.4.3). A significant, positive association was found between EFW>90th centile and caesarean section. The odds ratio was 3.1 (95% CI 2.26-4.2), the sensitivity was 0.45 (95% CI 0.39-0.51) and the specificity was 0.79 (95% CI 0.76-0.82). These resulted in a positive likelihood ratio of 2.15 and a negative likelihood ratio of 0.69.

Four studies that demonstrated positive and significant associations when the populations had a 'background rate of diabetes'. Chauhan (179) stated a background rate without giving the percentage. Levine (183) recorded the background rate of diabetes in the study population as 22%. Parry (182) recorded the background rate of diabetes in their study as 6-15% and Weiner (168) as a population with a mixed diabetes rate of 9.2%.

The study authored by Chauhan (179) was a retrospective case control study of 61 patients. The ultrasound report used was the most recent performed after 30weeks (>90% >36w). They recorded four separate test measures for each patient. An EFW≥4500g, an EFW≥4000-4999g, an AC≥90th centile and an AC≥35cm. The study was assessed to be of high

quality due to meeting five of the pre-specified quality criteria (Fig 2. Section 5.2.1).

There was a positive significant association between EFW ≥90th centile and caesarean section with an odds ratio of 2.9 (95% CI 1.04-8.0). The sensitivity was 0.49(95% CI 0.35-0.63) and the specificity was 0.75 (95% CI 0.59-0.91), giving a positive likelihood ratio of 1.96 and a negative likelihood ratio of 0.68.

There was also a significant association found between an AC≥35cm and caesarean section. The odds ratio was 3.1 (95% CI 1.18-8.2). The sensitivity was 0.63 (95% CI 0.50-0.77) and the specificity was 0.64 (95% CI 0.47-0.82). This resulted in a positive likelihood ratio of 1.77 and negative likelihood ratio of 0.57.

There was no significant association found between EFW≥4500g and caesarean; odds ratio 4.5 (95% CI 0.52-38.6), or for EFW≥4000g-4999g and caesarean; odds ratio 2.1 (95% CI 0.73-6.2).

Levine (183) authored a retrospective cohort study of 406 patients. Ultrasound was performed at greater than 36w gestation. The test measure used was EFW≥90th centile. The study was assessed to be of high quality due to meeting five of the quality criteria (Fig 2. Section 5.2.1). A significant association was found between EFW≥90th centile and caesarean section with an odds ratio of 2.4 (95% CI 1.41-4.1). The sensitivity of the

test was low, at 0.25 (95% CI 0.18-0.32) but the specificity was higher at 0.88 (95% CI 0.84-0.92). This resulted in a positive likelihood ratio of 2.05 and a negative likelihood ratio of 0.85.

The study authored by Parry (182) was a retrospective cohort study of 264 patients. The ultrasound was performed at greater than 37w gestation and the test measure used was an EFW≥4000g. The study met five of the predefined quality criteria and was therefore assessed to be of high quality (Fig 2. Section 5.2.1). There was a positive significant association found between EFW≥4000g and caesarean section. The odds ratio was 3.2 (95% CI 1.90-5.4) with a sensitivity of 0.68 (95% CI 0.59-0.77) and a specificity of 0.60 (95% CI 0.52-0.68). The positive likelihood ratio calculated as 1.70 and negative likelihood ratio as 0.53.

The prospective study by Weiner (168) had 315 participants (study detailed in 5.4.1). This study showed a positive and significant association between EFW≥4000g and caesarean section with an odds ratio of 3.1 (95% CI 1.93-5.0). The sensitivity was 0.60 (95% CI 0.51-0.69) and specificity was 0.67 (95% CI 0.61-0.74) leading to a positive likelihood ratio of 1.84 and a negative likelihood ratio of 0.59.

Two studies found positive, significant associations between test measures and caesarean section in populations that excluded patients with diabetes.

The large study by Melamed (177) with 1938 patients was of prospective case-control design (study detailed in 5.4.3). There was a significant positive association found between an EFW=4000-4999g and caesarean section. The odds ratio was 3.9 (95% CI 3.19-4.8), the sensitivity was 0.51 (95% CI 0.48-0.51) and the specificity was 0.79 (95% CI 0.76-0.81). The calculated positive likelihood ratio was 2.41and the negative likelihood ratio was 0.62.

Peleg (178) performed a retrospective cohort study of 443 patients (study detailed in 5.4.3). The results were presented by dividing into two groups, a group that excluded elective caesarean section and a group that included all emergency and elective caesareans as a total. There was a significant association found for EFW≥4kg and caesareans excluding elective caesarean with an odds ratio of 5.6 (95% CI 3.141-10.1). The sensitivity was 0.79 (95% CI 0.70-0.88) and specificity was 0.60 (95% CI 0.55-0.66). This resulted in a positive likelihood ratio of 1.98 and a negative likelihood ratio of 0.35. There was a greater significant association found for EFW≥4kg and the total number of caesareans. The odds ratio was 9.0 (95% CI 5.27-15.4) with a sensitivity of 0.86 (95% CI 0.80-0.92) and a specificity of 0.60 (95% CI 0.55-0.65). The positive likelihood ratio for this group was 2.14 and the negative likelihood ratio was 0.24.

Two studies recorded a positive and significant association in a population where the diabetes status was unknown.

The study, authored by Aviram (169) was a large retrospective cohort study of 3319 participants (study detailed in 5.4.3). There was a small, significant association found between a HC>90th centile (>347mm) and caesarean section. The odds ratio was 1.5 (95% CI 1.02-2.2), the sensitivity was 0.13 (95% CI 0.09-0.16) and the specificity was 0.91 (95% CI 0.90-0.92). The resultant positive likelihood ratio was 1.42 and the negative likelihood ratio was 0.90.

Tatarova (181) was the author of a large retrospective cohort of 1847 patients. Ultrasound was performed within seven days of delivery and the test measure of fetal macrosomia was an estimated fetal weight of greater than or equal to 4000g. The study included both elective and emergency caesarean sections together, as their outcome. The study was assessed to be of high quality due meeting five of the pre-specified criteria (Fig 2. Section 5.2.1) but please note that which five criteria were met for this study are not recorded in full on this figure as the physical translated paper was lost during the write-up process prior to creation of the figure). There was a strong and significant association found between an EFW≥4000g and caesarean section. The odds ratio was 9.2 (95% CI 5.78-14.6), the sensitivity was low at 0.16(95% CI 0.12-0.20) but the specificity was high at 0.98 (95% CI 0.97-0.99). This gave a positive likelihood ratio of 7.86 and a negative likelihood ratio of 0.85.

The systematic review found that the test measures in five studies (166, 170, 180, 181, 184, 185) revealed positive but non-significant associations with caesarean section.

The studies authored by Bochner (170) and Simpson (184) included patients with GDM and pre-existing diabetes whilst the study by Nelson (180, 186) just included patients with GDM.

The retrospective cohort study authored by Bochner (170) had 201 patients (study detailed in 5.4.3). There was no significant association found between AC>90th centile and caesarean section: Odds ratio 1.8 (95% CI 0.80-4.3).

The study by Simpson (184) was a retrospective cohort of 413 participants. The ultrasound was performed at between 28-32weeks gestation and the ultrasound measure for fetal macrosomia was an AC≥70th centile. The study met four of the pre-specified criteria and therefore was assessed to be of high quality (Fig 2. Section 5.2.1). There was no significant association found between an AC≥70th centile and caesarean section: Odds ratio 1.3 (95% CI 0.76-2.4).

Nelson (180) conducted a cohort study of 86 patients. Ultrasound was performed at 28-32+6w gestation. The test measure used was an EFW≥75th centile. The results were calculated for two groups of the cohort

patients. Firstly, in the subgroup who had caesarean section for labour arrest and, secondly, the total rate of caesarean section (including elective caesarean). The study was assessed to be of high quality due to meeting five of the pre-specified criteria (Fig 2. Section 5.2.1). There was a positive but not significant association between EFW>75th centile and caesarean section for labour arrest. The odds ratio was 3.7 (95% CI 0.91-14.8). There was no association found between EFW>75th centile and the total caesarean section rate: Odds ratio 1.4 (95% CI 0.59-3.3).

The final two studies that produced positive but non-significant associations with caesarean section had unknown diabetes status in their patient populations.

The study by Basit (166) was a retrospective cohort of 104 patients (study detailed in 5.4.3). There was no significant association found between EFW>95th centile and caesarean delivery: Odds ratio 1.1 (95% CI 0.44-2.9).

The study authored by Kehl (185) included 258 patients and was of a prospective cohort design. Ultrasound was performed within three days of delivery at a gestation of greater than 37 weeks. The test measure was an AC≥36cm. The study was assessed to be of high quality due to meeting seven of the pre-defined criteria (Fig 2. Section 5.2.1). There was not a

significant association found between AC≥36cm and caesarean section: Odds ratio 1.6 (95% CI 0.55-4.4).

Table 4 contains a summary of all the individual study systematic review results for caesarean section:

Table 4: Caesarean section. The odds ratio and 95% confidence intervals by ultrasound measurement

Study	Study size	Predicted macrosomia definition	Odds ratio	95% Confidence intervals
Parry, 2000	264	EFW>=4000g Hadlock	3.2	1.9 to 5.4
Chauhan, 2011	77	AC>= 35cm	3.1	1.2 to 8.2
Bochner, 1987	201	AC>90th centile	1.8	0.8 to 4.3
Chavan, 2013	296	AC>or=75th centile	1.7	1.0 to 3.0
Melamed, 2010	1938	EFW =4000-4999g	3.9	3.2 to 4.8
Chauhan, 2011	77	EFW>=4000g	2.1	0.7 to 6.2
Chauhan, 2011	77	EFW>=4500g	4.5	0.5 to 38.6
Chauhan, 2011	77	EFW>=90th centile	2.9	1.0 to 8.0
Levine, 1992	406	EFW>=90th centile Hadlock	2.4	1.4 to 4.1
Scifres, 2015	903	EFW>90th centile Hadlock	3.1	2.3 to 4.2
Basit, 2010	104	EFW>95th centile	1.1	0.4 to 2.9
Aviram, 2014	3319	HC>90th centile (>347mm)	1.5	1.0 to 2.2
Peleg, 2015	443	EFW>=4000g Hadlock	5.6	3.1 to 10.1 .
Nelson, 2009	86	EFW>75th centile	3.7	0.9 to 14.8
Kehl, 2011	258	AC>=36cm	1.6	0.5 to 4.4
Simpson, 2015	413	AC>or=70th centile	1.3	0.8 to 2.4
Tatarova, 2004	1847	EFW>=4000g	9.2	5.8 to 14.6
Weiner, 2002	315	EFW>=4000g Shepard	3.1	1.9 to 5.0
Nelson, 2009	86	EFW>75th centile	1.4	0.6 to 3.3
Peleg, 2015	443	EFW>=4000g Hadlock	9.0	5.3 to 15.4

5.5 Maternal outcomes

In addition to these primary outcomes, some studies assessed other different maternal outcomes. There were not enough studies testing the same test measure paired with maternal outcome to allow any further meta-analyses. Therefore, the following maternal outcome results are presented for each study individually.

5.5.1 Second/Third degree tear

Kehl (185) measured this outcome in a prospective cohort study. The study included 258 participants with unknown diabetes status (study detailed in 5.4.5). There was no association found between second and third degree tear and an AC≥36cm: Odds ratio 0.9 (95% CI 0.36-2.40).

5.5.2 Third/Fourth degree tear

Three studies measured the outcome of Third/Fourth degree tear. The studies by Melamed (177) and Peleg (178) excluded women with diabetes whilst the study by Scifres (176) included exclusively women with GDM. Each of these studies are summarised in turn below.

The study by Melamed (177) was a case control study with 1938 participants (study detailed in 5.4.3). There was no association found between EFW= 4000-4999g and Third/Fourth degree tear; odds ratio 0.1 (95% CI 0.01-2.10),

Peleg (178) published a retrospective cohort study of 443 participants (study detailed in 5.4.3). There was no significant association established between EFW≥ 4000g and 3rd/4th degree tear with an odds ratio of 2.30 (95% CI 0.38-13.90).

The study by Scifres was a retrospective cohort study. It included 903 participants, all with GDM (study detailed in 5.4.3). There was no association demonstrated between Third/Fourth degree tears and an EFW≥ 90th centile: Odds ratio 0.70 (0.22-2.0 95% CI).

5.5.3 Postpartum haemorrhage (PPH)

As previously summarised (study detailed in 5.4.3), the case control study by Melamed (177) had 1938 participants without diabetes. There was not a significant association found between PPH and EFW= 4000-4999g. The odds ratio was 1.5 (0.79-2.90 95% CI).

PPH was also assessed as an outcome in the 104 patient, retrospective cohort study by Basit (166). There was no association was found between PPH and EFW≥95th centile: Odds ratio 0.2 (95% CI 0.08- 0.60).

5.6 Fetal outcomes

In addition to our primary outcome measure of BPI, many studies assessed other different fetal outcomes, but there were not enough studies testing the same fetal outcomes to allow any further meta-analyses. Therefore,

these following fetal outcome results are calculated for each study individually. The incidence of all of these fetal complications is commoner with diabetes regardless of size (95) . All of these studies, except one, were conducted using maternal populations exclusively with diabetes. Eken (174) was the only study that had a "background rate" of diabetes.

5.6.1 Neonatal Intensive Care Unit (NICU) admission

One retrospective cohort study of 296 patients with gestational diabetes, authored by Chavan (173) used NICU admission as an outcome measure (study detailed in 5.4.3) There was a significant, increased occurrence of NICU admission when abdominal circumference was greater than or equal to the 75th centile; odds ratio 2.0 (95% CI 1.17-3.4). The sensitivity was 0.55 (95% CI 0.43-0.67) and specificity was 0.62 (95% CI 0.56-0.69) giving a positive likelihood ratio of 1.45 and a negative likelihood ratio of 0.72.

5.6.2 Neonatal hypoglycaemia

There were only studies in women with diabetes and it is known that this outcome can occur due to DM, independent of fetal size (187).

In the study by Chavan (study detailed in 5.4.5) (173) of 296 mothers with gestational diabetes a significant positive association was found between neonatal hypoglycaemia and an AC≥75th centile: Odds ratio 2.8 (95% CI 1.07-7.1). The sensitivity was 0.65 (95% CI 0.44-0.86) and the specificity was 0.6 (95% CI 0.54-0.66). These results calculated as a positive likelihood ratio of 1.62 and a negative likelihood ratio of 0.59.

One other study also looked at the outcome of neonatal hypoglycaemia. In this retrospective cohort study by Simpson (184) there were 413 mothers, all with gestational diabetes (study detailed in 5.4.5). They only studied neonatal hypoglycaemia in relation to an AC≥70th centile. There was no association found for these parameters: Odds ratio 1.0 (95% CI 0.67-1.7).

5.6.3 Neonatal hyperbilirubinaemia and APGAR <7

The aforementioned (study detailed in 5.4.3) Chavan (173) study of 296 patients with gestational diabetes also assessed two other fetal outcome measures in relation to AC≥ 75th centile. There was no association found with neonatal hyperbilirubinaemia: Odds ratio 0.6 (95% CI 0.25-1.3) or one minute fetal APGAR<7: Odds ratio 1.0 (95% CI 0.49-2.1). Neither of these two fetal outcomes were measured by any other studies or in a population without diabetes.

5.6.4 Birth trauma

Birth trauma (defined as fetal fracture, haematoma and soft tissue injury) as a fetal outcome was measured and compared to a single specific test measure in one retrospective cohort study of 201 gestational diabetes patients by Bochner (170) (study detailed in 5.4.3). This study revealed that birth trauma was not predicted by an AC≥90th centile. There was not a significant association: Odds ratio 3.1 (95% CI 0.84-11.1).

5.6.5 Neonatal trauma

Neonatal trauma was assessed as a fetal outcome in two studies. The study by Melamed (177) was a large prospective case-control study with 1938 patients, none of whom had diabetes (study detailed in 5.4.3). Neonatal trauma (defined as skull fracture, subdural/cerebral haematoma, intraventricular haemorrhage and facial nerve injury) was not found to be associated with an EFW=4000-4999g: Odds ratio 0.5 (95% CI 0.10-2.1).

Cohen's 1996 study (167) was a retrospective cohort of 31 diabetes patients (study detailed in 5.4.1). In this study, there was no significant association found between neonatal trauma (defined as Erb's palsy or humeral fracture) and AD-BPD≥2.6cm: Odds ratio 3.1 (95% CI 0.14-70.7).

5.6.6 Clavicle fracture

The small study by Eken (174), used a prospective cohort of 28 patients with fetal brachial plexus injury and recorded the incidence of clavicle fracture in association with three different, specific ultrasound measurements (study detailed in 5.4.3). As there was only one recorded case of clavicular fracture, the confidence intervals were huge and no associations could be established. For AC>350mm the odds ratio was 3.2 (95% CI 0.15-67). For AC>360mm the odds ratio was 0.2 (95% CI 0.007-3.2) and for EFW≥4000g the odds ratio was 0.5 (95% CI 0.02-11.2).

5.6.7 Neonatal composite

There was one study by Scifres (176) that used a neonatal composite. This composite was defined as hypoglycaemia, increased bilirubinaemia and the need for respiratory support at greater than 24 hours of life. This retrospective cohort of 903 patients with gestational diabetes (study detailed in 5.4.3) found that neonatal composite was not associated with an EFW≥90th centile: Odds ratio 0.9 (95% CI 0.59-1.3).

5.7 Subgroup analysis by maternal diabetes

There were six studies that used patient populations that all had diabetes; Bochner (170), Chavan (173), Cohen 1996 (167), Cohen 1999 (132) Elliott (135) and Nelson (186). Whilst, there were two studies by Melamed (177) and Peleg (178) that excluded all patients with diabetes.

All eight of these studies examined a different test measure so there were no comparable results for the diabetic and non-diabetic pregnancy populations. Therefore, it was not possible to complete meta-analysis for either of the populations, with or without diabetes, as a minimum of three trials with the same test measures and outcomes are required.

The individual study results for the diabetic and non-diabetic populations were categorised and presented within groups in the prior results sections for our three main outcomes of brachial plexus injury, shoulder dystocia and caesarean section.

6 Discussion

6.1 Main findings

This study shows that the ultrasound marker of fetal macrosomia most associated with prediction of shoulder dystocia is a difference in the abdominal diameter to the biparietal diameter of ≥ 2.6 cm. If this marker is positive then approximately 1 in 10 women will subsequently experience a shoulder dystocia. The more widely used estimated fetal weight > 4 kg has more uncertainty as a predictor of subsequent shoulder dystocia with only 7 in 100 women subsequently experiencing a shoulder dystocia.

The ultrasound marker of fetal macrosomia most associated with brachial plexus injury is an AC to HC difference of > 50 mm.

The ultrasound marker of fetal macrosomia most associated with caesarean delivery was an EFW > 4 kg. In my study if this marker was positive then approximately 1 in 2 women subsequently had a caesarean delivery. Due to the heterogeneity of the caesarean delivery results, I would not advise the use of this statistic in clinical practise.

6.2 Whole population

The quality assurance method used the STARD (162) and QUADAS (163) checklists that are validated for the reporting and methodological quality

of diagnostic test accuracy studies. These checklists instructed the previous validated technique that was used in similar studies (164, 165). When used in this systematic review this technique did not allow much differentiation between the quality of the included studies. This was illustrated by the fact that all but one of the 25 studies was classed as high quality. The included studies were demonstrated to be high quality observational studies but this does not necessarily mean that they were good for answering the question being asked. The best way to answer the question would have been to use RCTs. Observational studies are a lower standard of evidence than RCTs, which are more prone to bias and confounding, and cannot be used to show causality. However, there was no evidence found from randomised trials to evaluate which antenatal ultrasound measurements should be used once LGA is suspected (151). Therefore, the quality assessment used, although fairly non-discriminatory in this study, was the most appropriate choice. In retrospect, I have considered possible ways that I may have made the quality assessment more discriminatory. Two of the main factors associated with an increased level of bias are case-control studies as opposed to cohort and retrospective recruitment as opposed to prospective. Therefore, it may have been useful to stratify by awarding more quality points to studies that met the criteria of being a cohort or prospective study. Another consideration was to have more than the three study ratings of low (≤2 criteria), medium (3 criteria) and high (≥4 criteria). A fourth category of 'excellent' for studies meeting six or more criteria would have identified the studies by Cohen 1996 (167), Cohen 1999 (132), Kehl (185), Miller

(133), Rajan (134), Scifres (176) and Weiner (168) as being superior in quality to the other 18 studies.

The included studies originated from seven different countries with the vast majority of patients being from either the USA or Israel. Since the UK may have different antenatal, maternity and postnatal care from the USA and Israel, this review may have uncertain relevance to our society. The array of countries will mean that there are potentially different populations and standards of care between the studies. This will be most relevant when considering the caesarean section rate, which differs greatly between different countries (188). The most researched countries in our study were the USA with a caesarean section rate of 32% (189) and Israel's of 15% (190). This should be of particular relevance as elective caesarean was considered in the caesarean rate total in many of the studies (168, 176, 179, 181, 183, 184).

Diabetes was worth considering as a sub group because it is possible that the presence of maternal diabetes may affect the pathophysiology of fetal macrosomic growth and therefore there may be different ultrasound markers that identify adverse fetal outcome in populations with and without maternal diabetes. Although I analysed all the results for diabetes it was not possible to complete any meta-analysis. The six studies that exclusively used patients with diabetes and the two studies that excluded diabetes all measured different test measures for predicting fetal macrosomia. Therefore, comparison was not possible.

The timing of the ultrasound within the third trimester may also have affected the results. For my study, the third trimester was used but the studies seemed to fall into two subsections: early 3rd trimester (28-32w) and late 3rd trimester (36w-birth). It is possible that the ultrasound markers of fetal macrosomia that are associated with adverse outcome are different between these two subgroups. There were not enough studies to consider a sub-group analysis, four studies scanned at 28-32w and eighteen studies at >36w. For the outcome of shoulder dystocia, only one (25%) of the studies at 28-32w (Bochner (170)) found a significant positive association compared to five (36%) of the studies performed at >36w (Cohen 1999 (132), Chaabane (172), Miller (133), Rajan (134), Buikema (171)). It is likely that later timed ultrasound studies are more likely to detect fetal macrosomia correctly and therefore a more accurate test may be more likely to be associated with adverse outcomes. None of the studies looked at the outcome compared to two different time-point ultrasound measures, and therefore on the same patients, so a comparison was not possible.

The timing of the studies in history ranged from 1982-2015. Ultrasound technology advanced significantly in this period so that later ultrasound equipment will have been of better quality and therefore potentially more accurate measurements will have been gained in later dated studies. This systematic review did not differentiate between the different decades therefore this is a possible cause of inaccuracy. By 1982, real time

scanners, which had been produced in the 1970's had become inexpensive and widely available (191). The relative ease of use of the probe meant that many fetal structures were measured and charts created. Despite this, the standard measurements of BPD, HC and AC that were developed when US was static continued to be used for assessing fetal growth. The only addition at this time was FL, which was used in the Hadlock equations for fetal weight and growth predictions in 1985 (192). This evidence shows that the measurements have been consistent through the time scale of the systematic review but there is no doubt the technology has improved and more accurate images and measurements will have been gained over the years.

It is important to highlight that the RCT by Boulvain (5) that recommended induction for suspected fetal macrosomia to reduce the incidence of shoulder dystocia was published in 2015. Therefore, all of the trials found and included in this systematic review predate this publication. It is possible that there will have been altered clinical practices and therefore outcomes in the management of fetal macrosomia since this time.

6.3 Main outcomes: BPI, Shoulder dystocia, Caesarean section

6.3.1 Brachial plexus injury

Only a small number of studies recorded BPI as an outcome. Due to BPI being a rare adverse outcome, with an incidence of 0.42 per 1000 live

births (1 in 2300) (72), this made it a more difficult outcome to assess. The study by Cohen (167) was a very small, retrospective study of just 31 patients meaning that study was not large enough to have produced any significant results. Weiner, a prospective cohort study, was larger with 315 patients and had potentially more chance of an association gaining significance. The cohort, however, had only one BPI adverse outcome recorded also meaning that significance of any association found could not be attained. Although a positive association was found between AC to HC difference of greater than or equal to 50mm and BPI in the Endres study (131), there were no other studies or test measures to compare. The Endres paper only measured this one test measure so there were no comparisons within the study itself either. At most, our study can suggest that there may be a positive association between AC to HC difference≥50mm. This does not exclude that there may be other measures that may be associated with BPI. The power calculation to determine the minimum number of subjects for adequate study power of a diagnostic accuracy for BPI (with an incidence of 0.42%) calculates a sample size of 47000 (193). A very large study would therefore be needed to assess this outcome prospectively.

6.3.2 Shoulder dystocia

There were a large number of studies considering a shoulder dystocia outcome with large number of participants totalling 11447. Shoulder dystocia is also a more common outcome with an incidence of 0.58-0.70% (63) of all deliveries and 11% of macrosomic deliveries (defined as fetal

weight >4000g in this study) (64). There were therefore enough studies to perform meta-analysis.

The widely accepted recommendations are that elective caesarean section is only advised for fetal macrosomia if the baby's weight is estimated to be above 5kg in the absence of diabetes (119). and this is echoed in the ACOG guidelines (13). This recommendation was based on a decision analysis of the monetary costs of elective caesarean for different predicted birth weights. They estimated that 3695 caesarean deliveries need to be performed with suspected fetal macrosomia of >4500g to prevent a single case of permanent BPI, hence the higher calculated figure of 5kg being the EFW recommended to trigger elective caesarean section. In the UK the RCOG notes the Rouse and ACOG findings in their shoulder dystocia guideline (63) but does not provide any management recommendation for suspected macrosomia in non-diabetic women. It is accepted that elective caesarean will eliminate shoulder dystocia and BPI without increasing maternal mortality but because the US prediction of fetal macrosomia is inaccurate and BPIs are usually temporary, the justification for elective caesarean section fails on monetary and clinical grounds. That said, the ACOG (13) and RCOG (63) guidance for women with diabetes does recommend elective caesarean section for predicted fetal macrosomia of >4.5kg, showing that they do value the EFW measurement enough in this scenario. This is also due to the fact that in women with diabetes the shoulder dystocia risk is much higher and only 443 caesarean sections would need to be performed to prevent one case of permanent BPI at an

EFW>4.5kg; a much lower number needed to treat than in the non-diabetic population (number needed to treat of 3695(119)). All of these current recommendations are based around using EFW to predict fetal macrosomia. However, the meta-analysis result of this study suggested that there is not a significant association found between EFW>4kg (OR 1.97 (95% CI 0.90-4.28)) and shoulder dystocia. There was, however, a positive and significant association found between an AC>90th centile (OR 2.72 (95% CI 1.24-5.96)) and AD-BPD≥2.6cm (OR 5.27 (95% CI 2.85-9.77)) and shoulder dystocia.. This makes sense clinically as in diabetes, the babies are more likely to grow with disproportionate macrosomia, laying fat around their chest and abdomen rather than their head (98). Diabetes does not tend to affect the HC or FL measurements of a fetus and it therefore follows that AC should be a more sensitive measure of shoulder dystocia in pregnancies with diabetes. This may not be the case for the non-diabetic population.

The shoulder dystocia meta-analyses revealed two test measures of predicted fetal macrosomia as significant associations. The strengths of the meta-analysis were the large patient numbers and the fact that it enabled comparison between three different test measures. I used the most up to date and recommended techniques for performing and interpreting meta-analysis (194, 195). Each of the meta-analyses did not have the same population of patients so there may be bias in these results. Meta-analysis illustrated the most reliable comparison of this study due to

the larger patient numbers and the fact that the calculation process takes into account the size of each individual study.

This systematic review also revealed eight individual studies that produced positive associations between predictive tests of fetal macrosomia and shoulder dystocia. Four (50%) of these significant findings studies used compound test measurements that looked at the difference between the head and abdominal size. Three of these studies (38%) demonstrated the association between AC and shoulder dystocia and only one study (13%) showed a link between an EFW>90th centile and SD. These eight studies positive associations could not be accounted for by diabetes as the four compound measurement studies had populations of GDM, no diabetes and two studies with background levels of diabetes. The three AC associations were from three different populations, one with GDM, one with a background rate of diabetes and the third an unknown diabetes population. The one study that found a link between SD and EFW>90th centile had an unknown diabetes population. This array of findings was in keeping with the meta-analysis results suggesting that compound measurements are the best predictors, followed by AC measurement and EFW measurements were the least commonly associated.

The non-significant association study findings included six (40%) AC test measurements, five (33%) EFW measurements, two (17%) head circumference measurements and two (17%) compound measurements.

The compound measurements that performed so well in the significant positive results were an abdominal measurement (AC or AD) minus a head measurement (BPD or HC). One of the two studies with compound measurements that did not find significant results measured HC divided by AC. This was the only study that assessed a ratio. These findings suggest that a head measurement minus an abdominal measurement is a good predictive test but that a ratio of head to abdomen is not. The other negative compound measurement study that tested AD-BPD≥2.6cm was very small. There were only 31 patients, so despite a very high odds ratio of 10.3, this was unable to demonstrate significance due to type two error and therefore may well have reached significance if the patient numbers had been larger. This small study did not have the power to answer the question. None of the included observational studies documented calculating the power to decide their sample sizes.

When looking at the overall shoulder dystocia findings, nearly all of the studies had odds ratios greater than one but the fetal macrosomia measurements that looked at the comparison between the head and abdominal measurements in each baby were superior to either measurement performed individually. This makes sense when it is considered clinically. In shoulder dystocia there are two factors influencing the occurrence: the dimensions of the mother's pelvic outlet and the size of the fetal head and abdomen in comparison to the outlet. It seems logical that a compound measurement that takes into account the difference in how much bigger the abdomen is than the head may best predict shoulder

dystocia. When a smaller head fits through, there will be a larger body than the head causing the shoulder dystocia. The difference between the individual fetus's head and body are more important than the overall size or weight of the fetus. It is well known that almost half (48%) of shoulder dystocias occur in babies that are not deemed large, when the definition of macrosomia is 4kg at birth (63). What I believe has not been considered before in a systematic review is if these smaller than <4kg fetuses that experience shoulder dystocia had an AD-BPD≥2.6cm, a residual AD-BPD=0.33-2cm or an AC-HC≥50mm.

It is worth commenting on the results generated from the few studies that compared more than one test measure within the same population. The study by Miller (133) was of high quality with a lower risk of type two error due to being a cohort and study population size of 322. It demonstrated that AD-BPD>=2.6cm was significantly associated with shoulder dystocia with a high odds ratio of 7.2 (95% CI 2.89-17.7) and a significant, high positive likelihood ratio of 4.48. Within the same population there was no significant association demonstrated for AC>95th centile, EFW>97th centile or EFW>4000g. As all four measurements were made on the same population at the same time, within 2 weeks of birth, this suggests that a combined abdominal/head difference is better at predicting shoulder dystocia than AC or EFW alone. Rajan (134) measured two test measures on the same population and like the Miller study (133) found again that AD-BPD≥2.6cm was significantly associated with shoulder dystocia, with an odds ratio of 3.6 (95% CI 1.56-9.0) but that AC>90th centile had an

odds ratio of 1.8 (95% CI 0.69-4.5) but was not significant. This supports the conclusion that compound measurement is a better predictor of shoulder dystocia. The third study that measured more than one test measure on the same population was very small with just 31 patients. Eken was unable to show a significant association for AC>350mm, AC>360mm or EFW>4000g.

The case control study by Melamed is worth particular mention due to it being very large with 1938 patients and therefore having a low risk of type two error. It demonstrated that there was not a significant association between EFW=4000-4999g within 3 days of delivery. Due to the size of this study, this is a significant negative finding for EFW relating to shoulder dystocia.

I have heard opinions expressed as to whether the incidence of shoulder dystocia actually matters when there is no injury to the baby. I disagree with this opinion as the trauma of the experience of shoulder dystocia can have a profound effect on the mother, father and, or health care providers. This can result in post-traumatic stress disorder (PTSD) or other mental health difficulties and can have repercussions on the planning of future deliveries, either their existence or their mode of birth and the risks that can ensue from this (194). For healthcare practitioners it is known that difficult birth experiences can also cause PTSD symptoms (195) and affect their future clinical decision making (196) and this means shoulder dystocia can have a lasting impact for patients unrelated to the original

incident. There is no doubt that the incidence and management of shoulder dystocia causes distress and anxiety for future pregnancies. Thus, prevention of this complication of birth is important.

6.3.3 Caesarean section

There were a large number of studies (15) that recorded the outcome of caesarean section with a total of 11630 participants. Caesarean is a very common outcome so this resulted in enough studies to perform meta-analysis.

Despite the fact there were numerous studies, the only meta-analyses possible were for EFW≥4000g and EFW>90th centile. The first meta-analysis was good due to being large, as it included 4569 patients. There was a positive, significant association found between EFW≥4000g and caesarean section. The five studies that measured EFW≥4000g included three that had excluded elective caesarean sections (177, 178, 182) and two that included both elective caesarean and caesarean for labour dystocia (179, 181). The second meta-analysis demonstrated that there was a positive significant association between EFW≥90th centile and caesarean section. All three of the articles (179) included in this meta-analysis did not differentiate between elective caesarean and caesarean for labour dystocia. There is therefore potential bias in both of these meta-analyses results due to treatment paradox. The treatment paradox in these meta-analyses can be described as the prediction of fetal macrosomia defined by EFW ≥4000g or EFW ≥90th centile having caused clinicians to make an earlier decision for caesarean section when faced with managing

delay in labour. As both of these meta-analyses were testing EFW, this did not enable the comparison of any different test measures, particularly AC and compound measurements as in the shoulder dystocia data.

The individual study data when grouped into all diabetes, background diabetes, no diabetes or unknown diabetes status did not suggest any patterns of types of test measurement that were better than any others as there were different studies that found significant and non-significant associations for many of the same test measures.

Although not originally specified in my research I did retrospectively ascertain the studies that excluded and included elective caesareans. The studies that included caesarean for labour arrest were authored by Aviram (169), Basit (166), Bochner (170), Kehl (185), Melamed (177), Nelson (186), Parry (182) and Peleg (178). Out of these studies the positive significant associations in individual studies were found with the test measures EFW≥4000g (OR 3.2 and OR 5.6), EFW=4000-4999g (OR 3.9) and HC>90th centile (OR 1.5). There was no association found for EFW>75th centile, EFW>95th centile, AC≥36cm and AC>90th centile. This may suggest that a high EFW is a better predictor of caesarean section in labour than a raised AC. However, it could also indicate that clinicians are more likely to intervene when they perceive that the baby is large and have an earlier recourse to caesarean section than for appropriate for gestational age babies, as the results of the EFW are revealed to clinicians. The study by Bochner (170) was not a large study so it was difficult to reach significance but the results were suggestive of a possible link

between AC>90th centile at 30-33 weeks and caesarean section in women with GDM.

Unsurprisingly, as elective caesarean is often recommended or chosen for fetal macrosomia, six of the seven studies that included elective caesarean found a positive significant association with caesarean and the test measures. The seven studies that included elective caesarean section were by Chavan (173), Chauhan (179), Levine (183), Scifres (176), Simpson (184), Tatarova (181) and Weiner (168). The positive significant test measures were EFW≥90th centile (found in three studies), EFW≥4000g (two studies) and AC≥35cm. The only non-significant test measure for total caesarean section was an AC≥70th centile found in the study by Simpson (184). The likely reasons for this study's negative findings were the fact that a 70th centile is within the normal range and the ultrasound test measurement was performed between 28-32w when the final decision for caesarean section was made many weeks after this scan finding.

The following studies are worth considering separately for a number of different reasons.

The study by Chauhan (179), although a small study of 61 patients, was the only study that assessed multiple test measures. Despite the small size, this study did find significance in the positive association with caesarean section for AC≥35cm (OR 3.1) and EFW≥90th centile (OR 2.9).

The same population did not show a significant association for an EFW≥4500g or EFW=4000-4999g. This study did not differentiate between elective and emergency caesarean section hence does include the bias of clinical decision-making. Despite this, it could be potentially useful as an indicative tool when seeing someone at 36w and counselling them that if any of the US between 28-36w had been positive for the AC≥35cm or EFW≥90th centile test measures they were more likely to require or be recommended a caesarean delivery.

Kehl's study (185) was rated very highly in quality assurance (see Figure 2) and was less likely to be subject to bias because it was one of the few prospective cohort studies and it excluded elective caesareans section. Due to its prospective nature there were a smaller number of true positives than in the retrospective studies and it was therefore difficult to attain significant results. The odds ratio (1.6 (95% CI 0.55-4.4)) was suggestive of AC≥36cm possibly being associated with the need for emergency caesarean section although further studies would be required to ascertain this. Interestingly, an AC>35cm was found to have a higher sensitivity (i.e. lower false negative rate) than EFW for prediction of actual birth weight in the systematic review by Malin (144). Similarly, the study by Nelson (186) was high quality due to cohort design and separated results for caesarean for labour arrest and for total rate including electives. A positive association was found but was not significant. There were only 86 patients in this study of ultrasound at 28-32+6 weeks, so there is a possible predictive ability of this test measure of EFW>75th centile but a larger study would be needed to reach significance.

The Levine cohort study with 406 patients assessed the test measure EFW≥90th centile. As elective caesareans were included, the results of this study actually say more about whether the sonographic suspicion of fetal macrosomia at term affected obstetric management. Interestingly the article found a significant association with an odds ratio of 2.4 but states that the significance of the association was lost if the elective caesarean sections were excluded.

Peleg's medium sized study (175) of 443 patients divided the patients into two groups, one that excluded elective caesarean section and a group that contained all electives and emergency caesareans. A significant positive association was found between EFW≥4000g and caesarean (excluding electives) with an odds ratio of 5.6 (3.14-10.1). An even greater significant positive association was found when electives were included (OR 9.0 (5.27-15.4)). These results therefore revealed more about current labour management practices when macrosomia is suspected, than caesarean section rates because of the fetal size itself. The article indicated that the ultrasonic prediction of fetal macrosomia (defined as EFW≥4000g in this study) raises concerns to both the mother and the physician, leading to an unjustified, significant increase in the caesarean section rate. The physician concerns led to a high rate of elective caesarean section and an earlier diagnosis of labour abnormalities.

The Tatarova (181) study was a large 1847 patient cohort study assessing the test measure EFW≥4000g. Elective caesarean sections were included and a strong and significant association was found OR 9.2 (5.78-14.6). As this study was performed in Bulgaria, it is difficult to consider it as a comparison to the other studies because the rate of elective caesarean in Bulgaria is higher at 43% and I was unable to retrieve any information on Bulgarian obstetric practices and labour management.

The study authored by Melamed (177) was potentially less prone to bias as it was a large prospective case-control study. The 1938 patients reduce the chance of type two error and increases the chance of significant results. The actual purpose of this study was to establish the effect of false diagnosis of macrosomia on maternal/perinatal outcomes. They therefore compared a group that were predicted to have an EFW of 3500-3999g to a group predicted as EFW=4000-4499g. The data was collected and presented in a way that made it possible to extract the emergency caesarean section rate that occurred in these two groups. This concept introduced possible bias to these results. Another important factor increasing bias is the fact that the study was not blinded. The article discusses the fact that clinicians who suspected fetal macrosomia based on an ultrasound prediction were more likely to diagnose labour abnormalities and perform caesarean section. It is therefore unsurprising that a significant positive association (OR 3.9) was found between EFW=4000-4999g and caesarean section.

Another study with a similar aim was the article by Parry, a medium-sized 264 patient, retrospective cohort. The study aimed to investigate if the rate of caesarean section is associated with the US prediction of fetal macrosomia of EFW>4000g (as opposed to EFW=3000-3999g). They compared false negatives and positives with the actual birth weights. Their results demonstrated that the sonographic prediction of macrosomia increased the relative risk of caesarean section (RR2.07 (1.10-3.92)) and that incorrect prediction (false negatives) of EFW<4000g significantly reduced the caesarean section rate. We were able to extract the data to compare the outcomes for the predicted EFW>4000g and EFW=3000-3999g groups and there was a positive significant association found between EFW≥4000g and caesarean section (Odds ratio 3.2). This study, unlike the others, sub-analysed their results by removing the confounders of multiparity and diabetes and found that for the nulliparous and nondiabetic subgroups the results were still significant. This study illustrates clearly that in all studies, the rate of caesarean section when EFW>4000g is falsely increased by the management of the labour based on that finding. This fact is applicable to all of the caesarean section data, so this demonstrates that this section of our systematic review tells us more about the current worldwide management of suspected fetal macrosomia rather than the absolute risk conferred by the estimation of the elevated EFW centile. Essentially the caesarean section results show an observation of the management of suspected fetal macrosomia versus non-suspected.

6.4 Maternal outcomes

There was no meta-analysis possible; therefore, the only evidence found was for single study outcomes. Very few of the included studies recorded ultrasound measures of predicted fetal macrosomia with maternal outcomes. Four separate studies found no significant positive associations between perineal trauma ($2^{nd}/3^{rd}$ or $3^{rd}/4^{th}$ degree tears) and predicted fetal macrosomia using the test measure definitions of EFW>=4000g, EFW=4000g-4999g, EFW>90^{th} centile and AC>=36cm. Many confounding factors can affect the risk of genital tract lacerations except the predicted size of the fetus. Fetal macrosomia and shoulder dystocia are associated with protracted labour and both of these circumstances increase the need for assisted delivery, which in itself increases the risk of genital tract lacerations (197).

This systematic review found no associations for either EFW>95^{th} centile or EFW=4000-4999g and postpartum haemorrhage. The study that assessed EFW=4000-4999g by Melamed (177) was large with 1938 patients. The possibility of type 2 error was reduced by the large sample size for this relatively common complication. It is known that more large for gestational age pregnancies result in postpartum haemorrhage. The large retrospective study by Weissmann-Brenner (83) showed a statistically significant association between postpartum haemorrhage and a birth weight of greater than 4500g but not for birth weight greater than 4000g. The difference in the results of this study to our systematic review could be explained by the fact that ultrasound prediction of EFW is

inaccurate, as opposed to retrospective analysis of absolute birthweight. The other reason could be the fetal macrosomia definition used. The Melamed study (177) included all EFW from 4000g to 4999g, which included both of Weissmann-Brenner's (83) groups; the >4500g significant association group and the 4000-4500g non-significant association group. It is therefore entirely possible that had Melamed (177) looked at the >4500g group in isolation, that a positive, significant association may also have been found. The second study by Basit (166) included in the systematic review that observed postpartum haemorrhage only assessed an EFW>95th centile, which would have compared to the >4500g group of actual birth weight assessed in the Weissmann-Brenner study (83). This study was less reliable in quality assessment and only included patients who gave birth to babies with an actual birth weight of >5000g. The control group who had an EFW of less than the 95th centile were actually inaccurate in their prediction and therefore were a high-risk group for postpartum haemorrhage themselves. It is therefore unsurprising that there was no association found because all of the study population had babies that had a birth weight over 5000g and we know that postpartum haemorrhage is associated with babies born with a weight > 4500g (83). This study did not comment on the mode of delivery in respect to the outcome of postpartum haemorrhage so this is a possible confounder.

The fact that none of the included studies demonstrated significant associations with these maternal adverse outcomes may suggest a number of factors. It is possible that the prediction of fetal macrosomia is associated with adverse maternal outcomes, or that the many maternal

factors that influence the occurrence of genital trauma and PPH are more contributory. With such limited study data on this subject, it is also possible though that there is an association with predicted fetal macrosomia that has not yet been proven.

6.5 Fetal outcomes

I have already considered the adverse fetal outcome of BPI. Here I consider the other fetal outcomes. There was no meta-analysis possible, only single study outcomes. It is known that the incidence of all of these fetal complications are more common when the mother has diabetes (65, 95). All of the studies in this systematic review, except one, that had a background rate of diabetes, used populations that all had maternal diabetes. Therefore, the question being asked in our study was whether a fetus that is predicted by ultrasound to be macrosomic is significantly more likely to have adverse fetal outcomes, irrespective of their mother's diabetes status.

The study by Chavan (173) found two significant positive associations with an AC>75th centile in mothers with gestational diabetes. This test measure was found to be associated with the need for NICU admission and neonatal hypoglycaemia. It is well known that women with less well-controlled diabetes grow more macrosomic babies and that the more poorly controlled the diabetes is, the higher the chance of NICU admission or neonatal hypoglycaemia (87, 95). Our findings therefore support these facts. This study was presented as a poster publication so there was less data about their population and methods supplied. It is worth considering

that some units may electively admit babies born to mothers with diabetes to the NICU due to their risk of hypoglycaemia and this may (1) have been a confounder. Neonatal hypoglycaemia is more common in preterm babies (198), but as our systematic review only included studies where the populations delivered at term (>37w), this possible confounder was removed by the study methodology.

Interestingly, the study by Simpson (184) with 413 mothers, who also had GDM did not find a significant association with neonatal hypoglycaemia when an AC>70th centile was used. This study performed the ultrasound scan between 28-32 weeks so it is possible that the lack of association found was due to the fact that the test measure was performed too early and that later in the third trimester, many of the previously AC<70th centile could have been predicted as AC>70th centile and this could have biased the results. It is also important to consider that the test measure of AC>70th centile is in itself unlikely to be linked to pathology, as it is a measure that includes approximately a third of the fetal population.

None of the other fetal outcomes studied showed a significant positive association. Perhaps this could be due to the rarity of some of the outcomes e.g. clavicular fracture (one adverse outcome / 28 patients who all experienced BPI), neonatal trauma (10/164 and 11/1938), Apgar<7 at birth (33/296), and neonatal hyperbilirubinaemia (30/296). Even a neonatal composite that tried to combine a few outcomes to make the

outcome more common (137/903), did not show an association. None of these studies were powered for these adverse outcomes.

6.6 Subgroup analysis by maternal diabetes

There was not enough data published on the prediction of fetal macrosomia to predict adverse outcomes to analyse and compare the groups of patients with or without diabetes. This systematic review has illustrated that there is a paucity of evidence on this subject with reference to diabetes status and that this an area that requires further research.

6.7 Strengths and limitations of the thesis

The study followed a prospectively registered protocol (199), using detailed methodology, large patient populations and up to date statistical techniques for the meta-analysis (194, 195). We demonstrated both the prognostic association of ultrasound markers with adverse outcome (odds ratio) and their predictive ability (positive predictive value).

The small studies and smaller number of studies for certain ultrasound markers may be associated with results that are more uncertain. This is somewhat adjusted for by the meta-analysis process and the resultant 95 % confidence intervals of the odds ratios.

A comparison of the accuracy of different tests using studies that did not compare the tests within the same study population could be prone to confounding. Only four of the included studies compared multiple measures of suspected fetal macrosomia (EFW, AC, AD-BPD, and HC) within the same population. All four of these studies recorded outcome data for shoulder dystocia and just one of these studies recorded outcome data for BPI. In the case of the outcome shoulder dystocia, the individual study results proved consistent with the meta-analysis findings that used all of the available studies to compare the multiple tests, suggesting reliability. For the outcome of BPI there were significantly less studies and therefore no meta-analysis was possible. The individual study results compared different populations and thus the results presented for BPI are at a higher risk of confounding.

One weakness of the study is that the data on caesarean section did not differentiate between elective and emergency caesarean and therefore did not account for 'treatment paradox'.

6.8 Strengths and limitations compared to other published research

Previous studies have focused on the diagnostic accuracy of antenatal prediction of fetal macrosomia in comparison to absolute birth weights. Currently, no other studies have systematically reviewed our index tests in relation to delivery and neonatal outcomes. A Cochrane review published

in 2016 (151) asked a similar question. Culliney et al. focused on the benefits and harms associated with different combinations of surveillance methods for the suspected LGA fetus but they found no RCTs that assessed the effect of the antenatal fetal surveillance regimens on important health outcomes for the mother and baby. The absence of randomised controlled trials revealed an area where research is needed. Our study relied on observational data in the form of cohort and case-control studies. Observational data is prone to bias. The main problem with observational studies is the presence of confounders and selection bias (which RCTs manage to prevent through randomisation and blinding). In the absence of RCT data, systematic review and meta-analysis of the observational trials became the best option to obtain data on this question. In this scenario, meta-analysis of the observational studies was able to create a new hypothesis and was the only way to answer this question.

This is the first study to our knowledge to examine which antenatal imaging measures of fetal macrosomia best predict adverse maternal and neonatal outcomes. No other review has attempted to compare different definitions of suspected fetal macrosomia to inform clinical practice. This is more useful information for clinicians needing to counsel women with suspected fetal macrosomia.

6.9 Clinical Implications

When counselling women with suspected fetal macrosomia this review and meta-analysis would suggest that compound measurements that look at the difference between the size of the fetal head and abdomen are the best predictors of the likelihood of shoulder dystocia. These compound measurements are not being routinely calculated in our UK practise but AC is always measured. AC was also demonstrated by this research to be a more accurate test for predicting shoulder dystocia than the more conventionally used EFW.

It is evidenced (2, 3, 5, 6, 200) that induction of labour between 38 and 39 weeks for women carrying a fetus with predicted fetal macrosomia can decrease the incidence of shoulder dystocia without increasing the caesarean section rate. It is therefore important that this new research be considered in how we select the women to offer induction for fetal macrosomia. In the absence of RCT evidence on this subject, the observational findings of this systematic review provide the best available guidance for selecting the ultrasound measure of predicted fetal macrosomia. This research would suggest to first consider the difference in measurement of the fetal head and the abdomen, followed by the AC and not recommending caesarean for an EFW alone (4500g in diabetes and 5000g in non-diabetes pregnancies (13)). Therefore, a large for gestational age fetus with a symmetrical head and abdominal measurement would not routinely be offered induction of labour at 39 weeks. Conversely, a fetus predicted within the normal growth centiles (HC

and AC<90th centile) may have a large difference between AC and HC and therefore be offered induction of labour at 39 weeks to reduce the risk of shoulder dystocia and adverse outcome.

7 Conclusions

7.1 Summary

Estimated fetal weight is the most widely used ultrasound marker to predict fetal macrosomia in the UK. However, this study suggests that other markers have a higher positive predictive value for adverse outcomes associated with fetal macrosomia.

When counselling women with suspected fetal macrosomia and offering interventions to reduce shoulder dystocia, clinicians should be aware that this research would suggest that ultrasound markers that look at the difference between the size of the fetal abdomen and head or the size of the fetal abdomen alone could more accurately predict shoulder dystocia than estimated fetal weight.

7.2 Unanswered questions and future work

At the time of the study searches and inclusion there were other studies in process collecting ultrasound data and adverse outcome. The Pregnancy Outcome Prediction (POP) study (35) investigated the relationship between serial prenatal ultrasonography, biomarkers, placental phenotype and adverse outcome in nulliparous women. The prospective cohort study was sufficiently large enough to be powered for relatively uncommon adverse outcomes (4212 patients). The final ultrasound was performed at

36 weeks gestation, when EFW was calculated to detect SGA pregnancies to compare the diagnostic effectiveness of universal ultrasound as a screening test for SGA. Although the subject of excessive growth was not considered in the original POP study, other recent analyses of the study cohort led to publication that the maternal diagnosis of GDM at or after 28 weeks was preceded by excessive AC growth between 20w and 28w (201). The study did not provide any analysis or recommendation of the prediction of fetal macrosomia in the third trimester. This is an example of a study that will have recorded a predictor of fetal macrosomia at 36w and adverse birth outcomes. At the time, the research group was unwilling to share the data to become part of this systematic review. An IPD systematic review and meta-analysis would provide the gold standard of systematic reviews on this subject. To achieve this, the data from recent studies such as POP would need to be shared and collaboration required. The results of our systematic review have clearly illustrated that they may not be using the best predictors of fetal macrosomia for prediction of adverse outcome in these studies such as POP (that used EFW, AC>90^{th} centile and HC:AC ratio<10^{th} centile). The best next step forward would be to perform an IPD meta-analysis to investigate whether AD-BPD≥2.6cm, AC>90^{th} (or 97^{th}) centile or EFW>90^{th} (or 97^{th}) centile is most significantly associated with adverse outcome. It would be valuable to perform this IPD meta-analysis to instruct a test accuracy RCT that could provide higher-grade evidence of this association. There are no studies developed to answer the question that we have asked of the data. An IPD analysis done with available data would be asking a different question than that original data was collected to ask. This could affect the validity of the results. Advantages of the IPD

analysis include allowing adjustment for the same variables across studies e.g. for parity or diabetes, the fact that they may more often obtain unpublished data, the ability to assess the adequacy of randomisation, permitting data checking, data updating and checking of analyses. Although, for some of the older studies, the patient level data may not still be available for meta-analysis. The resources, time and cooperation required for such IPD studies can act as a barrier to their performance.

As we found no studies that were developed to actually ask the question that we asked, regarding whether any ultrasound measures of fetal macrosomia are associated with adverse outcomes, it would be important to consider if a prospective observational study to ask the question we have asked, needs doing prior to a RCT test accuracy study. It is a concern that the RCTs that have been completed and those in planning to evaluate induction of labour versus expectant management for fetal macrosomia have been using EFW as their test measure. Without better evidence of the test measure that best predicts adverse outcome, these trials may be using a suboptimal test measure to select inclusion of women for the RCT. My systematic review suggests that we could be using better screening tools for fetal macrosomia and therefore further research is needed urgently to instruct the RCTs for interventions aimed at reducing the adverse outcomes of fetal macrosomia.

8 References

1. Boulet SL, Alexander GR, Salihu HM, Pass M. Macrosomic births in the united states: determinants, outcomes, and proposed grades of risk. Am J Obstet Gynecol. 2003;188(5):1372-8.

2. Boulvain M, Irion O, Dowswell T, Thornton JG. Induction of labour at or near term for suspected fetal macrosomia. Cochrane Database Syst Rev. 2016(5):CD000938.

3. Gonen O, Rosen DJ, Dolfin Z, Tepper R, Markov S, Fejgin MD. Induction of labor versus expectant management in macrosomia: a randomized study. Obstet Gynecol. 1997;89(6):913-7.

4. Tey A, Eriksen N, Blanco J. A prospective randomized trial of induction versus expectant management in nondiabetic pregnancies with fetal macrosomia. American Journal of Obstetrics & Gynecology. 1995;172(1):293.

5. Boulvain M, Senat MV, Perrotin F, Winer N, Beucher G, Subtil D, et al. Induction of labour versus expectant management for large-for-date fetuses: a randomised controlled trial. Lancet. 2015;385(9987):2600-5.

6. Kean L. L*eave alone or induce for the big baby (LIBBY)*. 1998. National Research Register. www.update-software.com/NRR (accessed 10 July 2017).

7. Chauhan SP, Grobman WA, Gherman RA, Chauhan VB, Chang G, Magann EF, et al. Suspicion and treatment of the macrosomic fetus: a review. American Journal of Obstetrics & Gynecology. 2005;193(2):332-46.

8. Abramowicz J, Ahn J. Fetal macrosomia In: UpToDate, Post TW (Ed), UpToDate, Waltham, MA.: UpToDate; 2017.

9. Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, et al. International standards for fetal growth based on serial

ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Lancet. 2014;384(9946):869-79.

10. Delpapa EH, Mueller-Heubach E. Pregnancy outcome following ultrasound diagnosis of macrosomia. Obstetrics & Gynecology. 1991;78(3 Pt 1):340-3.

11. Chatfield J. ACOG issues guidelines on fetal macrosomia. American College of Obstetricians and Gynecologists. Am Fam Physician. 2001;64(1):169-70.

12. Ju H, Chadha Y, Donovan T, O'Rourke P. Fetal macrosomia and pregnancy outcomes. Aust N Z J Obstet Gynaecol. 2009;49(5):504-9.

13. ACOG. Fetal macrosomia. (Practice Bulletin No. 173). Obstet Gynecol [Internet]. 2016 (128):[e195-209 pp.].

14. Gardosi J. Customized fetal growth standards: rationale and clinical application. Semin Perinatol. 2004;28(1):33-40.

15. Mikolajczyk RT, Zhang J, Betran AP, Souza JP, Mori R, Gülmezoglu AM, et al. A global reference for fetal-weight and birthweight percentiles. Lancet. 2011;377(9780):1855-61.

16. Mohammadbeigi A, Farhadifar F, Soufi Zadeh N, Mohammadsalehi N, Rezaiee M, Aghaei M. Fetal macrosomia: risk factors, maternal, and perinatal outcome. Ann Med Health Sci Res. 2013;3(4):546-50.

17. Gaudet L, Ferraro ZM, Wen SW, Walker M. Maternal obesity and occurrence of fetal macrosomia: a systematic review and meta-analysis. Biomed Res Int. 2014;2014:640291.

18. Stock SJ, Myers J. Defining Abnormal Fetal Growth and Perinatal Risk: Population or Customized Standards? PLOS Medicine. 2017;14(1):e1002229.

19. Demerath EW, Choh AC, Czerwinski SA, Lee M, Sun SS, Chumlea WC, et al. Genetic and environmental influences on infant weight and weight change: the Fels Longitudinal Study. Am J Hum Biol. 2007;19(5):692-702.

20. Alexander GR, Kogan MD, Himes JH. 1994-1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. Matern Child Health J. 1999;3(4):225-31.

21. Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. Radiology. 1991;181(1):129-33.

22. Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. Customised antenatal growth charts. Lancet. 1992;339(8788):283-7.

23. Larkin J, Speer P, Simhan H. Use of a customized growth standard to identify LGA neonates predicts intrapartum morbidity. Reproductive Sciences. 2010;17 (3 SUPPL. 1) (pp 196A).

24. Larkin JC, Speer PD, Simhan HN. A customized standard of large size for gestational age to predict intrapartum morbidity. Am J Obstet Gynecol. 2011;204(6):499.e1-10.

25. Pasupathy D, McCowan LM, Poston L, Kenny LC, Dekker GA, North RA, et al. Perinatal outcomes in large infants using customised birthweight centiles and conventional measures of high birthweight. Paediatr Perinat Epidemiol. 2012;26(6):543-52.

26. Costantine MM, Mele L, Landon MB, Spong CY, Ramin SM, Casey B, et al. Customized versus population approach for evaluation of fetal overgrowth. Am J Perinatol. 2013;30(7):565-72.

27. Chiossi G, Pedroza C, Costantine MM, Truong VTT, Gargano G, Saade GR. Customized vs population-based growth charts to identify neonates at risk of adverse outcome: systematic review and Bayesian meta-analysis of observational studies. Ultrasound Obstet Gynecol. 2016.

28. Sebire NJ, Jolly M, Harris JP, Wadsworth J, Joffe M, Beard RW, et al. Maternal obesity and pregnancy outcome: a study of 287,213 pregnancies in London. Int J Obes Relat Metab Disord. 2001;25(8):1175-82.

29. O'Brien TE, Ray JG, Chan WS. Maternal body mass index and the risk of preeclampsia: a systematic overview. Epidemiology. 2003;14(3):368-74.

30. Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Neerup Jensen L, et al. The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. PLoS Med. 2017;14(1):e1002220.

31. Anderson NH, Sadler LC, McKinlay CJD, McCowan LME. INTERGROWTH-21st vs customized birthweight standards for identification of perinatal mortality and morbidity. Am J Obstet Gynecol. 2016;214(4):509.e1-.e7.

32. Francis A, Hugh O, Gardosi J. Customized vs INTERGROWTH-21. Am J Obstet Gynecol. 2018;218(2S):S692-S9.

33. Pritchard N, Lindquist A, Siqueira IDA, Walker SP, Permezel M. INTERGROWTH-21st compared with GROW customized centiles in the detection of adverse perinatal outcomes at term. J Matern Fetal Neonatal Med. 2020;33(6):961-6.

34. Kabiri D, Romero R, Gudicha DW, Hernandez-Andrade E, Pacora P, Benshalom-Tirosh N, et al. Prediction of adverse perinatal outcome by fetal biometry: comparison of customized and population-based standards. Ultrasound Obstet Gynecol. 2020;55(2):177-88.

35. Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. Lancet. 2015;386(10008):2089-97.

36. Araujo Júnior E, Peixoto AB, Zamarian AC, Elito Júnior J, Tonni G. Macrosomia. Best Pract Res Clin Obstet Gynaecol. 2017;38:83-96.

37. Jolly MC, Sebire NJ, Harris JP, Regan L, Robinson S. Risk factors for macrosomia and its clinical consequences: a study of 350,311 pregnancies. Eur J Obstet Gynecol Reprod Biol. 2003;111(1):9-14.

38. Henriksen T. The macrosomic fetus: a challenge in current obstetrics. Acta Obstet Gynecol Scand. 2008;87(2):134-45.

39. Koyanagi A, Zhang J, Dagvadorj A, Hirayama F, Shibuya K, Souza JP, et al. Macrosomia in 23 developing countries: an analysis of a multicountry, facility-based, cross-sectional survey. Lancet. 2013;381(9865):476-83.

40. Lu Y, Zhang J, Lu X, Xi W, Li Z. Secular trends of macrosomia in southeast China, 1994-2005. BMC Public Health. 2011;11:818.

41. Roglic G. Diabetes in women: the global perspective. Int J Gynaecol Obstet. 2009;104 Suppl 1:S11-3.

42. Martorell R, Khan LK, Hughes ML, Grummer-Strawn LM. Obesity in women from developing countries. Eur J Clin Nutr. 2000;54(3):247-52.

43. Adesina OA, Olayemi O. Fetal macrosomia at the University College Hospital, Ibadan: a 3-year review. J Obstet Gynaecol. 2003;23(1):30-3.

44. Najmi RS. Distribution of birthweights of hospital born Pakistani infants. J Pak Med Assoc. 2000;50(4):121-4.

45. Serirat S, Deerochanawong C, Sunthornthepvarakul T, Jinayon P. Gestational diabetes mellitus. J Med Assoc Thai. 1992;75(6):315-9.

46. Chen CP, Chang FM, Chang CH, Lin YS, Chou CY, Ko HC. Prediction of fetal macrosomia by single ultrasonic fetal biometry. Journal of the Formosan Medical Association. 1993;92(1):24-8.

47. Ørskou J, Kesmodel U, Henriksen TB, Secher NJ. An increasing proportion of infants weigh more than 4000 grams at birth. Acta Obstet Gynecol Scand. 2001;80(10):931-6.

48. Mikulandra F, Stojnić E, Perisa M, Merlak I, Sikić D, Zenić N. Fetal macrosomia--pregnancy and delivery. Zentralbl Gynakol. 1993;115(12):553-61.

49. Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet. 2014;384(9946):857-68.

50. Dennedy MC, Dunne F. Macrosomia: defining the problem worldwide. Lancet. 2013;381(9865):435-6.

51. Cresswell JA, Campbell OM, De Silva MJ, Filippi V. Effect of maternal obesity on neonatal death in sub-Saharan Africa: multivariable analysis of 27 national datasets. Lancet. 2012;380(9850):1325-30.

52. O'Sullivan EP, Avalos G, O'Reilly M, Dennedy MC, Gaffney G, Dunne F, et al. Atlantic Diabetes in Pregnancy (DIP): the prevalence and outcomes of gestational diabetes mellitus using new diagnostic criteria. Diabetologia. 2011;54(7):1670-5.

53. Sørensen HT, Sabroe S, Rothman KJ, Gillman M, Fischer P, Sørensen TI. Relation between weight and length at birth and body mass index in young adulthood: cohort study. BMJ. 1997;315(7116):1137.

54. Seidman DS, Laor A, Gale R, Stevenson DK, Danon YL. A longitudinal study of birth weight and being overweight in late adolescence. Am J Dis Child. 1991;145(7):782-5.

55. Chiavaroli V, Giannini C, D'Adamo E, de Giorgis T, Chiarelli F, Mohn A. Insulin resistance and oxidative stress in children born small and large for gestational age. Pediatrics. 2009;124(2):695-702.

56. Darendeliler F, Poyrazoglu S, Sancakli O, Bas F, Gokcay G, Aki S, et al. Adiponectin is an indicator of insulin resistance in non-obese prepubertal children born large for gestational age (LGA) and is affected by birth weight. Clin Endocrinol (Oxf). 2009;70(5):710-6.

57. Giapros V, Evagelidou E, Challa A, Kiortsis D, Drougia A, Andronikou S. Serum adiponectin and leptin levels and insulin resistance in children born large for gestational age are affected by the degree of overweight. Clin Endocrinol (Oxf). 2007;66(3):353-9.

58. Evagelidou EN, Kiortsis DN, Bairaktari ET, Giapros VI, Cholevas VK, Tzallas CS, et al. Lipid profile, glucose homeostasis, blood pressure, and obesity-anthropometric markers in macrosomic offspring of nondiabetic mothers. Diabetes Care. 2006;29(6):1197-201.

59. Mocanu EV, Greene RA, Byrne BM, Turner MJ. Obstetric and neonatal outcome of babies weighing more than 4.5 kg: an analysis by parity. Eur J Obstet Gynecol Reprod Biol. 2000;92(2):229-33.

60. Sultan AH, Kamm MA, Hudson CN, Bartram CI. Third degree obstetric anal sphincter tears: risk factors and outcome of primary repair. BMJ. 1994;308(6933):887-91.

61. Nesbitt TS, Gilbert WM, Herrchen B. Shoulder dystocia and associated risk factors with macrosomic infants born in California. Am J Obstet Gynecol. 1998;179(2):476-80.

62. NHSLA. Ten Years of Maternity Claims: An Analysis of NHS Litigation Authority Data2012 27/7/17. Available from:

http://www.nhsla.com/Safety/Documents/Ten%20Years%20of%20Maternity%20Claims%20-%20An%20Analysis%20of%20the%20NHS%20LA%20Data%20-%20October%202012.pdf.

63. Shoulder Dystocia: Green Top Guideline No.42 [Internet]. RCOG. 2012 [cited 22nd February 2017]. Available from: https://www.rcog.org.uk/globalassets/documents/guidelines/gtg_42.pdf.

64. Gregory KD, Henry OA, Ramicone E, Chan LS, Platt LD. Maternal and infant complications in high and normal weight infants by method of delivery. Obstet Gynecol. 1998;92(4 Pt 1):507-13.

65. Tsur A, Sergienko R, Wiznitzer A, Zlotnik A, Sheiner E. Critical analysis of risk factors for shoulder dystocia. Arch Gynecol Obstet. 2012;285(5):1225-9.

66. Resnik R. Management of shoulder girdle dystocia. Clin Obstet Gynecol. 1980;23(2):559-64.

67. Spong CY, Beall M, Rodrigues D, Ross MG. An objective definition of shoulder dystocia: prolonged head-to-body delivery intervals and/or the use of ancillary obstetric maneuvers. Obstet Gynecol. 1995;86(3):433-6.

68. Beall MH, Spong C, McKay J, Ross MG. Objective definition of shoulder dystocia: a prospective evaluation. Am J Obstet Gynecol. 1998;179(4):934-7.

69. Baskett TF, Allen AC. Perinatal implications of shoulder dystocia. Obstet Gynecol. 1995;86(1):14-7.

70. Acker DB, Sachs BP, Friedman EA. Risk factors for shoulder dystocia. Obstetrics & Gynecology. 1985;66(6):762-8.

71. Dildy GA, Clark SL. Shoulder dystocia: risk identification. Clin Obstet Gynecol. 2000;43(2):265-82.

72. Evans-Jones G, Kay SP, Weindling AM, Cranny G, Ward A, Bradshaw A, et al. Congenital brachial palsy: incidence, causes, and outcome in the United

Kingdom and Republic of Ireland. Arch Dis Child Fetal Neonatal Ed. 2003;88(3):F185-9.

73. Gherman RB, Ouzounian JG, Satin AJ, Goodwin TM, Phelan JP. A comparison of shoulder dystocia-associated transient and permanent brachial plexus palsies. Obstet Gynecol. 2003;102(3):544-8.

74. Lagerkvist AL, Johansson U, Johansson A, Bager B, Uvebrant P. Obstetric brachial plexus palsy: a prospective, population-based study of incidence, recovery, and residual impairment at 18 months of age. Dev Med Child Neurol. 2010;52(6):529-34.

75. Dyachenko A, Ciampi A, Fahey J, Mighty H, Oppenheimer L, Hamilton EF. Prediction of risk for shoulder dystocia with neonatal injury. Am J Obstet Gynecol. 2006;195(6):1544-9.

76. Leung TY, Stuart O, Sahota DS, Suen SS, Lau TK, Lao TT. Head-to-body delivery interval and risk of fetal acidosis and hypoxic ischaemic encephalopathy in shoulder dystocia: a retrospective review. BJOG. 2011;118(4):474-9.

77. Siggelkow W, Boehm D, Skala C, Grosslercher M, Schmidt M, Koelbl H. The influence of macrosomia on the duration of labor, the mode of delivery and intrapartum complications. Arch Gynecol Obstet. 2008;278(6):547-53.

78. Bjørstad AR, Irgens-Hansen K, Daltveit AK, Irgens LM. Macrosomia: mode of delivery and pregnancy outcome. Acta Obstet Gynecol Scand. 2010;89(5):664-9.

79. Konje JC, Ladipo OA. Nutrition and obstructed labor. Am J Clin Nutr. 2000;72(1 Suppl):291S-7S.

80. Herbst MA. Treatment of suspected fetal macrosomia: a cost-effectiveness analysis. American Journal of Obstetrics & Gynecology. 2005;193(3 Pt 2):1035-9.

81. Bais JM, Eskes M, Pel M, Bonsel GJ, Bleker OP. Postpartum haemorrhage in nulliparous women: incidence and risk factors in low and high risk women. A Dutch population-based cohort study on standard (> or = 500 ml) and severe (> or = 1000 ml) postpartum haemorrhage. Eur J Obstet Gynecol Reprod Biol. 2004;115(2):166-72.

82. Sosa CG, Althabe F, Belizán JM, Buekens P. Risk factors for postpartum hemorrhage in vaginal deliveries in a Latin-American population. Obstet Gynecol. 2009;113(6):1313-9.

83. Weissmann-Brenner A, Simchen MJ, Zilberberg E, Kalter A, Weisz B, Achiron R, et al. Maternal and neonatal outcomes of macrosomic pregnancies. Med Sci Monit. 2012;18(9):PH77-81.

84. Alsammani MA, Ahmed SR. Fetal and maternal outcomes in pregnancies complicated with fetal macrosomia. N Am J Med Sci. 2012;4(6):283-6.

85. RCOG. Prevention and Management of Postpartum Haemorrhage: Green-top Guideline No. 52. BJOG. 2017;124(5):e106-e49.

86. WHO. WHO recommendations for the prevention and treatment of postpartum haemorrhage. Italy: World Health Organisation; 2012.

87. Groenendaal F, Elferink-Stinkens PM, Netherlands Perinatal Registry. Hypoglycaemia and seizures in large-for-gestational-age (LGA) full-term neonates. Acta Paediatr. 2006;95(7):874-6.

88. Xu H, Simonet F, Luo ZC. Optimal birth weight percentile cut-offs in defining small- or large-for-gestational-age. Acta Paediatr. 2010;99(4):550-5.

89. Martin AO, Simpson JL, Ober C, Freinkel N. Frequency of diabetes mellitus in mothers of probands with gestational diabetes: possible maternal influence on the predisposition to gestational diabetes. Am J Obstet Gynecol. 1985;151(4):471-5.

90. Lin XH, Wu DD, Gao L, Zhang JY, Pan HT, Wang H, et al. Altered DNA methylation in neonates born large-for-gestational-age is associated with cardiometabolic risk in children. Oncotarget. 2016;7(52):86511-21.

91. Walker JD. NICE guidance on diabetes in pregnancy: management of diabetes and its complications from preconception to the postnatal period. NICE clinical guideline 3. London, February 2015. Diabet Med. 2015;25(9):1025-7.

92. Bell R, Glinianaia SV, Tennant PW, Bilous RW, Rankin J. Peri-conception hyperglycaemia and nephropathy are associated with risk of congenital anomaly in women with pre-existing diabetes: a population-based cohort study. Diabetologia. 2012;55:936–947

93. Al-Agha R, Firth RG, Byrne M, Murray S, Daly S, Foley M, et al. Outcome of pregnancy in type 1 diabetes mellitus (T1DMP): results from combined diabetes-obstetrical clinics in Dublin in three university teaching hospitals (1995-2006). Ir J Med Sci. 2012;181(1):105-9.

94. Balsells M, García-Patterson A, Gich I, Corcoy R. Maternal and fetal outcome in women with type 2 versus type 1 diabetes mellitus: a systematic review and metaanalysis. J Clin Endocrinol Metab. 2009;94(11):4284-91.

95. Esakoff TF, Cheng YW, Sparks TN, Caughey AB. The association between birthweight 4000 g or greater and perinatal outcomes in patients with and without gestational diabetes mellitus. Am J Obstet Gynecol. 2009;200(6):672.e1-4.

96. Gewolb IH. Effect of high glucose on fetal lung maturation at different times in gestation. Exp Lung Res. 1996;22(2):201-11.

97. Persson M, Pasupathy D, Hanson U, Norman M. Birth size distribution in 3,705 infants born to mothers with type 1 diabetes: a population-based study. Diabetes Care. 2011;34(5):1145-9.

98. Persson M, Fadl H, Hanson U, Pasupathy D. Disproportionate body composition and neonatal outcome in offspring of mothers with and without gestational diabetes mellitus. Diabetes Care. 2013;36(11):3543-8.

99. Weindling AM. The confidential enquiry into maternal and child health (CEMACH). Arch Dis Child. 2003;88(12):1034-7.

100. Kapoor N, Sankaran S, Hyer S, Shehata H. Diabetes in pregnancy: a review of current evidence. Curr Opin Obstet Gynecol. 2007;19(6):586-90.

101. McFarland MB, Trylovich CG, Langer O. Anthropometric differences in macrosomic infants of diabetic and nondiabetic mothers. J Matern Fetal Med. 1998;7(6):292-5.

102. Wong SF, Chan FY, Cincotta RB, Oats JJ, McIntyre HD. Sonographic estimation of fetal weight in macrosomic fetuses: diabetic versus non-diabetic pregnancies. Aust N Z J Obstet Gynaecol. 2001;41(4):429-32.

103. Benacerraf BR, Gelman R, Frigoletto FD. Sonographically estimated fetal weights: accuracy and limitation. Am J Obstet Gynecol. 1988;159(5):1118-21.

104. Rowan JA, Hague WM, Gao W, Battin MR, Moore MP, Investigators MT. Metformin versus insulin for the treatment of gestational diabetes. N Engl J Med. 2008;358(19):2003-15.

105. Chauhan SP, Sullivan CA, Lutton TC, Magann EF, Morrison JC. Parous patients' estimate of birth weight in postterm pregnancy. J Perinatol. 1995;15(3):192-4.

106. Ashrafganjooei T, Naderi T, Eshrati B, Babapoor N. Accuracy of ultrasound, clinical and maternal estimates of birth weight in term women. East Mediterr Health J. 2010;16(3):313-7.

107. Chauhan SP, West DJ, Scardo JA, Boyd JM, Joiner J, Hendrix NW. Antepartum detection of macrosomic fetus: clinical versus sonographic, including soft-tissue measurements. Obstetrics & Gynecology. 2000;95(5):639-42.

108. Goetzinger KR, Odibo AO, Shanks AL, Roehl KA, Cahill AG. Clinical accuracy of estimated fetal weight in term pregnancies in a teaching hospital. Journal of Maternal-Fetal & Neonatal Medicine. 2014;27(1):89-93.

109. 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. Obstetrics & Gynecology. 1998;91(1):72-7.

110. Gonen R, Spiegel D, Abend M. Is macrosomia predictable, and are shoulder dystocia and birth trauma preventable? Obstetrics and Gynecology. 1996;88(4):526-9.

111. Combs CA, Rosenn B, Miodovnik M, Siddiqi TA. Sonographic EFW and macrosomia: is there an optimum formula to predict diabetic fetal macrosomia? Journal of Maternal-Fetal Medicine. 2000;9(1):55-61.

112. Melamed N, Yogev Y, Meizner I, Mashiach R, Bardin R, Ben-Haroush A. Sonographic fetal weight estimation: which model should be used? J Ultrasound Med. 2009;28(5):617-29.

113. Rosati P, Arduini M, Giri C, Guariglia L. Ultrasonographic weight estimation in large for gestational age fetuses: a comparison of 17 sonographic formulas and four models algorithms. Journal of Maternal-Fetal & Neonatal Medicine. 2010;23(7):675-80.

114. 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(4):324-9.

115. Coomarasamy A, Connock M, Thornton J, Khan KS. Accuracy of ultrasound biometry in the prediction of macrosomia: a systematic quantitative review. BJOG: An International Journal of Obstetrics & Gynaecology. 2005;112(11):1461-6.

116. Faschingbauer F, Voigt F, Goecke TW, Siemer J, Beckmann MW, Yazdi B, et al. Fetal weight estimation in extreme macrosomia (> 4,500 g): comparison of 10 formulas. Ultraschall in der Medizin. 2012;33(7):E62-7.

117. Pinette MG, Pan Y, Pinette SG, Blackstone J, Garrett J, Cartin A. Estimation of fetal weight: mean value from multiple formulas. J Ultrasound Med. 1999;18(12):813-7.

118. Hoopmann M, Abele H, Wagner N, Wallwiener D, Kagan KO. Performance of 36 different weight estimation formulae in fetuses with macrosomia. Fetal Diagn Ther. 2010;27(4):204-13.

119. Rouse DJ, Owen J, Goldenberg RL, Cliver SP. The effectiveness and costs of elective cesarean delivery for fetal macrosomia diagnosed by ultrasound. JAMA. 1996;276(18):1480-6.

120. Gupta M, Hockley C, Quigley MA, Yeh P, Impey L. Antenatal and intrapartum prediction of shoulder dystocia. Eur J Obstet Gynecol Reprod Biol. 2010;151(2):134-9.

121. Porter B, Neely C, Szychowski J, Owen J. Ultrasonographic Fetal Weight Estimation: Should Macrosomia-Specific Formulas Be Utilized? American Journal of Perinatology. 2015;32(10):968-72.

122. Catalano PM, Tyzbir ED, Allen SR, McBean JH, McAuliffe TL. Evaluation of fetal growth by estimation of neonatal body composition. Obstet Gynecol. 1992;79(1):46-50.

123. Mintz MC, Landon MB, Gabbe SG, Marinelli DL, Ludmir J, Grumbach K, et al. Shoulder soft tissue width as a predictor of macrosomia in diabetic pregnancies. American Journal of Perinatology. 1989;6(2):240-3.

124. Sood AK, Yancey M, Richards D. Prediction of fetal macrosomia using humeral soft tissue thickness. Obstet Gynecol. 1995;85(6):937-40.

125. Landon MB, Sonek J, Foy P, Hamilton L, Gabbe SG. Sonographic measurement of fetal humeral soft tissue thickness in pregnancy complicated by GDM. Diabetes. 1991;40 Suppl 2:66-70.

126. Rigano S, Ferrazzi E, Radaelli T, Cetin ET, Pardi G. Sonographic measurements of subcutaneous fetal fat in pregnancies complicated by gestational diabetes and in normal pregnancies. Croat Med J. 2000;41(3):240-4.

127. 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(7):378-82.

128. Abramowicz JS, Robischon K, Cox C. Incorporating sonographic cheek-to-cheek diameter, biparietal diameter and abdominal circumference improves weight estimation in the macrosomic fetus. Ultrasound in Obstetrics & Gynecology. 1997;9(6):409-13.

129. Rossi AC, Vimercati A, Greco P, Baldassarra PF, Lestingi D, Laforgia N, et al. [Echographic measurement of subcutaneous adipose tissue as fetal growth index]. Acta Biomed Ateneo Parmense. 2000;71 Suppl 1:379-82.

130. Bailis A, Syzmanski L, Ibrahim S, Allen R, Gurewitsch E. Accelerated abdominal circumference growth (>75th%ile) in average gestational age fetuses increases risk of shoulder dystocia. Reproductive Sciences. 2010; 17 (3 SUPPL. 1) (pp 187A).

131. Endres L, Defranco E, Conyac T, Adams M, Zhou Y, Magner K, et al. Association of fetal abdominal-head circumference size difference with shoulder dystocia: A multicenter study. AJP Reports. 2015;5(2):e99-e104.

132. Cohen BF, Penning S, Ansley D, Porto M, Garite T. The incidence and severity of shoulder dystocia correlates with a sonographic measurement of asymmetry in patients with diabetes. Am J Perinatol. 1999;16(4):197-201.

133. Miller RS, Devine PC, Johnson EB. Sonographic fetal asymmetry predicts shoulder dystocia. Journal of Ultrasound in Medicine. 2007;26(11):1523-8.

134. Rajan PV, Chung JH, Porto M, Wing DA. Correlation of increased fetal asymmetry with shoulder dystocia in the nondiabetic woman with suspected macrosomia. J Reprod Med. 2009;54(8):478-82.

135. Elliott JP, Garite TJ, Freeman RK, McQuown DS, Patel JM. Ultrasonic prediction of fetal macrosomia in diabetic patients. Obstet Gynecol. 1982;60(2):159-62.

136. Nardozza LMM, Araujo Junior E, Vieira MF, Rolo LC, Moron AF. Estimate of birth weight using two - and three-dimensional ultrasonography. Revista da Associacao Medica Brasileira. 2010;56(2):204-8.

137. Nardozza LM, Vieira MF, Araujo Júnior E, 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(5):393-8.

138. 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(4):425-33.

139. Pagani G, Palai N, Zatti S, Fratelli N, Prefumo F, Frusca T. Fetal weight estimation in gestational diabetic pregnancies: comparison between conventional

and three-dimensional fractional thigh volume methods using gestation-adjusted projection. Ultrasound in Obstetrics & Gynecology. 2014;43(1):72-6.

140. Bennini JR, Marussi EF, Barini R, Faro C, Peralta CF. Birth-weight prediction by two- and three-dimensional ultrasound imaging. Ultrasound in Obstetrics & Gynecology. 2010;35(4):426-33.

141. Lindell G, Marsál K. Sonographic fetal weight estimation in prolonged pregnancy: comparative study of two- and three-dimensional methods. Ultrasound Obstet Gynecol. 2009;33(3):295-300.

142. Schild RL, Fimmers R, Hansmann M. Fetal weight estimation by three-dimensional ultrasound. Ultrasound in Obstetrics & Gynecology. 2000;16(5):445-52.

143. Hasenoehrl G, Pohlhammer A, Gruber R, Staudach A, Steiner H. Fetal weight estimation by 2D and 3D ultrasound: comparison of six formulas. Ultraschall in der Medizin. 2009;30(6):585-90.

144. 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: An International Journal of Obstetrics & Gynaecology. 2016;123(1):77-88.

145. Duncan KR. Fetal and placental volumetric and functional analysis using echo-planar imaging. Top Magn Reson Imaging. 2001;12(1):52-66.

146. Walsh JM, McAuliffe FM. Prediction and prevention of the macrosomic fetus. European Journal of Obstetrics, Gynecology, & Reproductive Biology. 2012;162(2):125-30.

147. Chiswick C, Reynolds RM, Denison F, Drake AJ, Forbes S, Newby DE, et al. Effect of metformin on maternal and fetal outcomes in obese pregnant women

(EMPOWaR): a randomised, double-blind, placebo-controlled trial. Lancet Diabetes Endocrinol. 2015;3(10):778-86.

148. Poston L, Bell R, Croker H, Flynn AC, Godfrey KM, Goff L, et al. Effect of a behavioural intervention in obese pregnant women (the UPBEAT study): a multicentre, randomised controlled trial. Lancet Diabetes Endocrinol. 2015;3(10):767-77.

149. Boulvain M, Senat MV, Perrotin F, Winer N, Beucher G, Subtil D, et al. Induction of labour versus expectant management for large-for-date fetuses: a randomised controlled trial. Lancet. 2015;385(9987):2600-5.

150. Quenby S GJ. Induction of labour for predicted macrosomia - The 'Big Baby Trial' [Available from: https://warwick.ac.uk/fac/sci/med/research/ctu/trials/bigbaby/.

151. Culliney Katherine AT, Parry Graham K, Brown J, Crowther Caroline A. Regimens of fetal surveillance of suspected large-for-gestational-age fetuses for improving health outcomes. Cochrane Database of Systematic Reviews [Internet]. 2016; (4). Available from: http://onlinelibrary.wiley.com/store/10.1002/14651858.CD011739.pub2/asset/CD011739.pdf?v=1&t=iry0dg5p&s=d3b23f10e609c5e55079449b5d14adfd3e51e9be.

152. Agoritsas T, Vandvik P, Neumann I. Finding Current Best Evidence. In: Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd Edition ed2015.

153. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med. 1997;126(5):376-80.

154. Deeks J, Wisniewski S, Davenport C. Chapter 4: Guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol. Cochrane Handbook for Systematic

Reviews of Diagnostic Test Accuracy: The Cochrane Collaboration; 2013. Available from: http://srdta.cochrane.org/.

155. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med. 1994;120(8):667-76.

156. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Eval Health Prof. 2002;25(1):76-97.

157. CRD UoY. Systematic reviews : CRD's guidance for undertaking reviews in health care. University of York, York: Centre for Reviews and Dissemination; 2009.

158. Khan KS, Dinnes J, Kleijnen J. Systematic reviews to evaluate diagnostic tests. Eur J Obstet Gynecol Reprod Biol. 2001;95(1):6-11.

159. Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. Stat Methods Med Res. 2017;26(4):1896-911.

160. Riley RD, Ahmed I, Debray TP, Willis BH, Noordzij JP, Higgins JP, et al. Summarising and validating test accuracy results across multiple studies for use in clinical practice. Stat Med. 2015;34(13):2081-103.

161. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol. 2006;6:9.

162. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.

163. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.

164. Malin GL, Morris RK, Riley R, Teune MJ, Khan KS. When is birthweight at term abnormally low? A systematic review and meta-analysis of the association and predictive ability of current birthweight standards for neonatal outcomes. BJOG. 2014;121(5):515-26.

165. Malin GL, Morris RK, Riley RD, Teune MJ, Khan KS. When is birthweight at term (≥37 weeks' gestation) abnormally low? A systematic review and meta-analysis of the prognostic and predictive ability of current birthweight standards for childhood and adult outcomes. BJOG. 2015;122(5):634-42.

166. Basit I, Ciprike V, Butler M, Daly S, Geary M. The antenatal and peripartum management ofpregnancies with macrosomic babies weighing >5000 g at two tertiary hospitals: A dublin experience. Archives of Disease in Childhood: Fetal and Neonatal Edition. 2010; 95 (pp Fa104).

167. Cohen B, Penning S, Major C, Ansley D, Porto M, Garite T. Sonographic prediction of shoulder dystocia in infants of diabetic mothers. Obstet Gynecol. 1996;88(1):10-3.

168. Weiner Z, Ben-Shlomo I, Beck-Fruchter R, Goldberg Y, Shalev E. Clinical and ultrasonographic weight estimation in large for gestational age fetus. Eur J Obstet Gynecol Reprod Biol. 2002;105(1):20-4.

169. Aviram A, Bardin R, Hadar E, Wiznitzer A, Yogev Y. The association between head circumference (HC)> 90th percentile at term and pregnancy outcome. American Journal of Obstetrics and Gynecology. 2014; 210 (1 SUPPL. 1) (pp S70).

170. Bochner CJ, Medearis AL, Williams IJ. Early third-trimester ultrasound screening in gestational diabetes to determine the risk of macrosomia and labor

dystocia at term. American Journal of Obstetrics and Gynecology. 1987;157(3):703-8.

171. Buikema T, Murphy J, Kazzi G. Fetal abdominal circumference (AC) of 35cm or greater predicts shoulder dystocia in fetuses presumed to have estimated fetal weight (EFW) appropriate for gestational age (AGA). American Journal of Obstetrics and Gynecology. 2014; 210 (1 SUPPL. 1) (pp S119).

172. Chaabane K, Trigui K, Louati D, Kebaili S, Gassara H, Dammak A, et al. Antenatal macrosomia prediction using sonographic fetal abdominal circumference in South Tunisia. The Pan African medical journal. 2013;14:111.

173. Chavan N, Jadhav A, Chun H, Monteagudo A. Utility of fetal abdominal circumference (AC) in predicting disease progression and perinatal outcomes in gestational diabetes mellitus (GDM). American Journal of Obstetrics and Gynecology. 2013; 208 (1 SUPPL.1) (pp S129-S130).

174. Eken M, Cinar M, Senol T, Ozkaya E, Karateke A. Six-year incidence and some features of cases of brachial plexus injury in a tertiary referral center. Turk Jinekoloji ve Obstetrik Dernegi Dergisi. 2015;12(2):71-4.

175. Peleg D, Warsof S, Wolf MF, Perlitz Y, Shachar IB. The effect of fetal weight estimation on cesarean section rates of the macrosomic neonate. American Journal of Obstetrics and Gynecology. 2014; 210 (1 SUPPL. 1) (pp S330).

176. Scifres CM, Feghali M, Dumont T, Althouse AD, Speer P, Caritis SN, et al. Large-for-Gestational-Age Ultrasound Diagnosis and Risk for Cesarean Delivery in Women With Gestational Diabetes Mellitus. Obstet Gynecol. 2015;126(5):978-86.

177. 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(2):225-30.

178. Peleg D, Warsof S, Wolf MF, Perlitz Y, Shachar IB. Counseling for fetal macrosomia: an estimated fetal weight of 4,000 g is excessively low. Am J Perinatol. 2015;32(1):71-4.

179. Chauhan S, Dattel B, Sbar W, Hendrix N, Weiner S. Detection of shoulder dystocia among newborns with birth weight of 4,500 g or more: Seven years experience. American Journal of Obstetrics and Gynecology. 2011; 204 (1 SUPPL.) (pp S149-S150).

180. Nelson L, Wharton B, Grobman W. Prediction of fetal macrosomia in term infants of diabetic mothers based on early third trimester ultrasound. American Journal of Obstetrics and Gynecology. 2009; 201 (6 SUPPL. 1) (pp S135).

181. Tatarova S, Popov I, Khristova P. [Fetal macrosomia: mode of delivery]. Akusherstvo i Ginekologiia. 2004;43(6):9-12.

182. 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(1):17-22.

183. 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? Obstetrics & Gynecology. 1992;79(1):55-8.

184. Simpson K, Pavicic M, Lee G. Does an enlarged abdominal circumference predict morbidity for patients diagnosed with gestational diabetes (GDM) using the IADPSG criteria? International Journal of Gynecology and Obstetrics. 2015; 131 (pp E195).

185. Kehl S, Brade J, Schmidt U, Berlit S, Bohlmann MK, Sutterlin M, et al. Role of fetal abdominal circumference as a prognostic parameter of perinatal complications. Arch Gynecol Obstet. 2011;284(6):1345-9.

186. Nelson L, Wharton B, Grobman WA. Prediction of large for gestational age birth weights in diabetic mothers based on early third-trimester sonography. Journal of Ultrasound in Medicine. 2011;30(12):1625-8.

187. Murphy HR. Continuous glucose monitoring targets in type 1 diabetes pregnancy: every 5% time in range matters. Diabetologia. 2019;62(7):1123-8.

188. Boerma T, Ronsmans C, Melesse DY, Barros AJD, Barros FC, Juan L, et al. Global epidemiology of use of and disparities in caesarean sections. Lancet. 2018;392(10155):1341-8.

189. Martin JA, Hamilton BE, Osterman MJK. Births in the United States, 2018. NCHS Data Brief. 2019(346):1-8.

190. Caesarean sections (indicator). [Internet]. [cited 20 March 2020].

191. Campbell S. A short history of sonography in obstetrics and gynaecology. Facts Views Vis Obgyn. 2013;5(3):213-29.

192. 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(2):503-5.

193. ClinCalc.com: Sample Size Calculator [Available from: https://clincalc.com/stats/SampleSize.aspx.

194. Beck CT. Middle Range Theory of Traumatic Childbirth: The Ever-Widening Ripple Effect. Glob Qual Nurs Res. 2015;2:2333393615575313.

195. Robins K. Psychological trauma of nurse-midwives following shoulder dystocia complicated by neonatal morbidity or mortality. 2012. Scholar Archive. 922. http://digitalcommons.ohsu.edu/etd/922

196. Minooee S, Cummins A, Sims DJ, Foureur M, Travaglia J. Scoping review of the impact of birth trauma on clinical decisions of midwives. J Eval Clin Pract. 2019.

197. Mazouni C, Menard JP, Porcu G, Cohen-Solal E, Heckenroth H, Gamerre M, et al. Maternal morbidity associated with obstetrical maneuvers in shoulder dystocia. Eur J Obstet Gynecol Reprod Biol. 2006;129(1):15-8.

198. Sharma A, Davis A, Shekhawat PS. Hypoglycemia in the preterm neonate: etiopathogenesis, diagnosis, management and long-term outcomes. Transl Pediatr. 2017;6(4):335-48.

199. Rebecca Robinson, Nia Wyn Jones. Which diagnostic definition of predicted fetal macrosomia best associates with adverse delivery and neonatal outcomes? A systematic review and meta-analysis. PROSPERO 2016 CRD42016046850 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42016046850

200. *Tey A*, *Eriksen N*, *Blanco J*. A prospective randomized trial of induction versus expectant management in nondiabetic pregnancies with fetal macrosomia. American Journal of Obstetrics & Gynecology. 1995;172(1):293.

201. Sovio U, Murphy HR, Smith GCS. Accelerated fetal growth prior to diagnosis of gestational diabetes mellitus: A prospective cohort study of nulliparous women. Diabetes Care. 2016;39(6):982-7.

9 Appendices

9.1 Appendix 1: PROSPERO Registration

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for Health Research

PROSPERO International prospective register of systematic reviews

Which diagnostic definition of predicted fetal macrosomia best associates with adverse delivery and neonatal outcomes? A systematic review and meta-analysis

Rebecca Robinson, Nia Wyn Jones

Citation

Rebecca Robinson, Nia Wyn Jones. Which diagnostic definition of predicted fetal macrosomia best associates with adverse delivery and neonatal outcomes? A systematic review and meta-analysis. PROSPERO 2016:CRD42016046850 Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016046850

Review question(s)

Which measurement of predicted fetal macrosomia most accurately predicts adverse delivery and neonatal outcome?

Searches

We will search the following electronic databases: MEDLINE, PubMed, EMBASE, The Cochrane Library (the Cochrane Central Register of Controlled Trials (CENTRAL)), Open Grey and the Web of Science.
The search strategy includes terms related to the population, index test and reference standard.
There will be no language restrictions.
The search will include articles published from database inception to July 2016.
Hand searching of major journals and reference checking will also be performed.
The searches will be re-run just before the final analyses and further studies retrieved for inclusion.

Types of study to be included

Inclusion: Test accuracy or epidemiological studies which consist of observational studies of defined populations. Cohort and case control studies will be used.
Exclusion: Case series of 10 or fewer participants.

Condition or domain being studied

Fetal macrosomia (many definitions; e.g. >4.5kg, >90th centile at birth) is associated with adverse delivery and neonatal outcomes. These adverse outcomes include shoulder dystocia (delay in the birth of the baby following the delivery of the head due to the shoulders impacting on the maternal pelvis), neonatal trauma, maternal perineal tears and postpartum haemorrhage.

Participants/ population

Population: Women with a singleton pregnancy.
Inclusion: Pregnant women who have had sonographic or MRI fetal measurements antenatally that have been recorded as predicted fetal macrosomia;
AND
That have had outcome measures of their term delivery (=37 completed weeks) also recorded.
Exclusion: Twin pregnancy or preterm delivery.

Intervention(s), exposure(s)

Index test: Any 2D or 3D ultrasound scan or magnetic resonance image (MRI) measurements defined as fetal macrosomia; including:
Estimated fetal weight = >4kg, 4-4.5kg, >4.5kg, >90th centile, >95th centile;
Abdominal circumference = >35cm, >36cm, >75th, >90th or > 95th centile;
Ratios of measurements and novel distances - e.g. abdominal diameter/biparietal diameter, femur length/abdominal circumference, cheek to cheek, humeral soft tissue thickness.

Comparator(s)/ control

Any reference standard looking at adverse delivery or neonatal outcome; including:
Shoulder dystocia;
Brachial plexus injury;
Caesarean section rate;
3rd or 4th degree perineal
trauma; Post-partum
haemorrhage.

Outcome(s)

Primary outcomes
To establish which measurement of predicted fetal macrosomia most accurately predicts adverse delivery and neonatal outcome.
Antenatal measurements from ultrasound or MRI in the 3rd trimester to detect fetal macrosomia.
Outcomes recorded at term birth - end of neonatal period (<28 days).
Secondary outcomes
To establish at what gestation time the scan measurement for predicted fetal macrosomia most accurately predicts adverse delivery and neonatal outcome.
The gestation at the time of the scan prediction of fetal macrosomia in the analysis will be considered, e.g. 28-32 weeks, 32-36 weeks, 36 weeks - term birth.

Data extraction, (selection and coding)

Studies will be selected by reviewing the electronic search citations, hand searching of the major journals and reference checking of review articles and studies selected for inclusion.
Data will be extracted using a standardised form, from the included studies, for assessment of study quality and evidence synthesis. The extracted information will include the methodology and timing of the index tests and reference standards that allow the generation of 2 x 2 tables (true positives, false positives, false negatives and true negatives) to calculate an estimate of the association between our index tests and reference standard outcomes.
The study selection and data extraction will be performed in duplicate by two members of the review team.

Risk of bias (quality) assessment

The studies will be assessed using the QUADAS-2 checklist.

Strategy for data synthesis

We will use aggregate data to quantitatively synthesise odds ratios (ORs) and 95% confidence intervals (95% CIs) for each index test-outcome pair. The results will be pooled for each index test using a random effect meta-analysis to reveal the potential prognostic association with outcome. If a strong and statistically significant prognostic association is identified between a test and an outcome measure, we will calculate sensitivity, specificity and

likelihood ratios from the 2 x 2 tables and then use a bivariate random-effects meta-analysis model to assess the predictive ability of a test.

Analysis of subgroups or subsets

If the necessary data are available subgroup analyses will be performed for gestation subsets at the time of the index test according to which index test performed was performed, the outcome measure and the study quality.

Dissemination plans
Publication through conference abstract and peer-reviewed journal.

Contact details for further information

Rebecca Robinson
School of Medicine, Department of Obstetrics and Gynaecology, D floor, East block, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Derby Road, Nottingham, NG7 2UH
drrebeccarobinson@gmail.com

Organisational affiliation of the review
None

Review team

Dr Rebecca Robinson, Nottingham University Hospitals NHS Trust
Dr Nia Wyn Jones, University of Nottingham

Collaborators
Mr George Bugg, Nottingham University Hospitals NHS Trust

Anticipated or actual start date 01 July 2016
Anticipated completion date 01 July 2017 **Funding sources/sponsors**
No external funding source.

Conflicts of interest
None known

Language
English

Country
England

Subject index terms status
Subject indexing assigned by CRD

Subject index terms

Caesarean Section; Diagnostic Techniques, Obstetrical and Gynaecological; Female; Fetal Macrosomia; Humans; Infant Health; Pregnancy; Pregnancy Complications; Pregnancy Outcome; Ultrasonography, Prenatal

Stage of review
Ongoing

Date of registration in PROSPERO 11 October 2016

Date of publication of this revision11 October 2016

Stage of review at time of this submission	Started	Completed
Preliminary searches	No	Yes
Piloting of the study selection process	Yes	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

9.2 Appendix 2: Search strategies for systematic review of macrosomia and outcome

9.2.1 Medline search strategy

MEDLINE_OVID_SEARCH 11/7/16	
1.	birth weight/
2.	birth weight.mp.
3.	birth-weight.mp.
4.	birthweight.mp.
5.	large-for-gestational-age.mp.
6.	large for gestational age.mp.
7.	large for date*.mp.
8.	large for gestation*.mp.
9.	lga.mp.
10.	Infant, large for gestational age/
11.	fetal macrosomia.mp.
12.	Infant, large for gestational age/
13.	fetal macrosomia.mp.
14.	macrosom*.mp.
15.	exp Fetal Macrosomia/
16.	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
17.	estimated fetal weight.mp.
18.	EFW.mp.
19.	abdominal circumference.mp.
20.	head circumference.mp.
21.	hadlock.mp.
22.	body constitution/ or "body weights and measures"/
23.	17 or 18 or 19 or 20 or 21 or 22
24.	shoulder dystocia.mp.
25.	brachial plexus injur*.mp.
26.	fracture*.mp.
27.	Fetal Monitoring/ or Infant, Newborn/ or Asphyxia Neonatorum/ or Cerebral Palsy/ or Asphyxia/ or Fetal Blood/ or Fetal Hypoxia/ or Delivery, Obstetric/ or Pregnancy/ or fetal asphyxia.mp. or Pregnancy Complications/
28.	caesarean section rate.mp.
29.	cesarean section/
30.	exp Asphyxia Neonatorum/
31.	Hypoxia-Ischemia, Brain/ and Infant, Newborn/
32.	pals*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
33.	24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
34.	16 and 23 and 33

Embase 1980-2016: the Medline Ovid search strategy was used for this database.

9.2.2 PubMed search strategy

Final PubMed search 13/7/16

Search	Add to builder	Query	Items found	Time
#122	Add	Search **(#119 and #120 and #121)**	1946	08:27:07
#121	Add	Search **(#95 or #116)**	70445	08:24:46
#120	Add	Search **(#103 or #115)**	324043	08:23:59
#119	Add	Search **(#113 or #117 or #118)**	382779	08:23:29
#118	Add	Search **caesarean section[MeSH Major Topic]**	20587	08:22:15
#117	Add	Search **infant, newborn[MeSH Major Topic]**	53190	08:21:51
#116	Add	Search **fetal macrosomia[MeSH Major Topic]**	1199	08:21:31
#115	Add	Search **ultrasonography, prenatal[MeSH Major Topic]**	16177	08:19:34
#114	Add	Search **(#95 and #103 and #113)**	279	08:15:29
#113	Add	Search **(#104 or #105 or #106 or #107 or #108 or #109 or #110 or #110 or #111 or #112)**	316680	08:14:46
#112	Add	Search **erb's pals*[Text Word]**	171	08:13:15
#111	Add	Search **hypoxia[Text Word]**	101265	08:12:56
#110	Add	Search **cerebral palsy[Text Word]**	22318	08:12:36
#109	Add	Search **asphyxia neonatorum[Text Word]** Schema: **syn**	7213	08:12:05
#108	Add	Search **caesarean section[Text Word]**	13723	08:11:47
#107	Add	Search **fracture[Text Word]**	162122	08:11:16
#106	Add	Search **brachial plexus injury[Text Word]**	1262	08:10:51
#105	Add	Search **brachial plexus[Text Word]**	12936	08:10:33
#104	Add	Search **shoulder dystocia[Text Word]**	1141	08:10:15
#103	Add	Search **(#96 or #97 or #98 or #99 or #100 or #101 or #102)**	308700	08:09:42
#102	Add	Search **body measures[Text Word]**	160	08:08:53
#101	Add	Search **body weight[Text Word]**	294110	08:08:32
#100	Add	Search **body constitution[Text Word]**	10546	08:07:35
#99	Add	Search **hadlock[Text Word]**	150	08:07:16
#98	Add	Search **head circumference[Text Word]**	5360	08:06:56
#97	Add	Search **abdominal circumference[Text Word]**	1822	08:06:39
#96	Add	Search **estimated fetal weight[Text Word]**	925	08:06:25
#95	Add	Search **(#93 not #94)**	70342	08:05:56
#94	Add	Search **small[Text Word]**	1166216	08:05:37
#93	Add	Search **(#87 or #88 or #89 or #90 or #91 or #92)**	79463	08:04:23
#92	Add	Search **macrosom*[Text Word]**	4011	08:01:51
#91	Add	Search **infant large for gestation*[Text Word]**	1067	08:01:11
#90	Add	Search **large for gestation*[Text Word]**	1539	08:00:42
#89	Add	Search **large for gestational age[Text Word]**	1506	08:00:21
#88	Add	Search **birth weight[Text Word]**	76782	07:59:49
#87	Add	Search **fetal macrosomia[Text Word]**	2369	07:59:07
#86	Add	Search **(#85 not #78)**	28281	07:58:13
#85	Add	Search **(#82 and #83 and #84)**	32870	07:49:31

Search	Add to builder	Query	Items found	Time
#84	Add	Search **(#79 or #35 or #53)**	969070	07:48:44
#83	Add	Search **(#68 or #81)**	747223	07:46:32
#82	Add	Search **(#67 or #10)**	96252	07:45:59
#81	Add	Search **"Ultrasonography, Prenatal"[Mesh]**	27557	07:44:26
#79	Add	Search **(#69 or #70 or #71 or #72 or #73 or #74 or #75 or #76 or #77)**	469631	07:27:42
#78	Add	Search **small**	1196208	07:26:16
#77	Add	Search **erb's pals***	175	07:25:49
#76	Add	Search **hypoxia**	131329	07:25:28
#75	Add	Search **cerebral palsy**	24756	07:25:14
#74	Add	Search **asphyxia neonatorum**	7223	07:24:50
#73	Add	Search **caesarean section**	53228	07:24:32
#72	Add	Search **fracture**	232154	07:24:13
#70	Add	Search **brachial plexus injury**	9879	07:24:01
#71	Add	Search **brachial plexus**	27188	07:23:54
#69	Add	Search **shoulder dystocia**	1217	07:23:09
#68	Add	Search **(#60 or #61 or #62 #63 or #64 or #65 or #66)**	721964	07:22:38
#67	Add	Search **(#3 or #4 or #56 or #57 or #58 or #59)**	96252	07:21:49
#66	Add	Search **body measures**	521382	07:20:45
#65	Add	Search **body weight**	529771	07:20:35
#64	Add	Search **body constitution**	532036	07:20:22
#63	Add	Search **hadlock**	389	07:20:05
#62	Add	Search **head circumference**	6220	07:19:54
#61	Add	Search **abdominal circumference**	7082	07:19:18
#60	Add	Search **estimated fetal weight**	2510	07:19:04
#59	Add	Search **macrosom***	4011	07:18:35
#58	Add	Search **infant large for gestation***	1067	07:18:10
#57	Add	Search **large for gestation***	1539	07:17:51
#56	Add	Search **large for gestational age**	6617	07:17:11
#4	Add	Search **fetal macrosomia**	2926	07:16:56
#3	Add	Search **birth weight**	91220	07:16:32
#55	Add	Search **(#13 and #25 and #54)**	26516	07:14:10
#54	Add	Search **(#27 or #29 or #31 or #33 or #35 or #37 or #41 or #43 or #46 or #49 or #51 or #53)**	776905	07:13:18
#53	Add	Search **"Cesarean Section"[Mesh]**	38166	07:06:18
#51	Add	Search **"Delivery, Obstetric"[Mesh]**	67928	07:05:50
#49	Add	Search **"Fetal Hypoxia"[Mesh]**	2913	07:05:27
#47	Add	Search **fetal hypoxia**	7091	07:05:08
#46	Add	Search **"Fetal Blood"[Mesh]**	27390	07:04:54
#44	Add	Search **fetal blood**	95047	07:04:41
#43	Add	Search **"Cerebral Palsy"[Mesh]**	17228	07:04:32
#41	Add	Search **"Asphyxia"[Mesh]**	5641	07:04:09

Search	Add to builder	Query	Items found	Time
#38	Add	Search **"Asphyxia**	3	07:03:50
#39	Add	Search **asphyxia**	17588	07:03:50
#37	Add	Search **"Asphyxia Neonatorum"[Mesh]**	7160	07:03:45
#35	Add	Search **"Infant, Newborn"[Mesh]**	533826	07:03:21
#33	Add	Search **"Fetal Monitoring"[Mesh] AND "Cardiotocography"[Mesh]**	1676	07:02:42
#31	Add	Search **"Fractures, Bone"[Mesh]**	157119	07:02:10
#29	Add	Search **"Birth Injuries"[Mesh]**	5210	07:01:36
#27	Add			

9.2.3 Cochrane library search strategy

Performed July 14th 2016

1# MeSH descriptor: [Fetal Macrosomia] explode all tree

#2 MeSH descriptor: [Ultrasonography, Prenatal] explode all trees

#3 MeSH descriptor: [Infant, Newborn] explode all trees

#4 MeSH descriptor: [Cesarean Section] explode all trees

#5 #3 or #4

#6 #1 and #2 and #5

#7 birth weight or birth-weight or birthweight

#8 large for gest* or large for date*

#9 fetal macrosom*

#10 #1 or #7 or #8

#11 estimated fetal weight

#12 abdominal circumference

#13 head circumference

#14 hadlock

#15 body constitution

#16 #2 or #11 or #12 or #13 or #14 or #15

#17 shoulder dystocia

#18 brachial plexus injur*

#19 fracture

#20 fetal monitoring

#21 caesarean section

#22 asphyxia neonatorum

#23 fetal hypoxia

#24 erb's pals*

#25 brachial plexus neuropath*

#26 #3 or #4 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25

#27 small

#28 #26 not #27

#29 #10 and #16 and #28

9.3 Appendix 3: Characteristics of the included studies

Table 5: Characteristics of the included studies

Primary reference	Sample size	Study design	Gestation at time of scan	Diabetes status	Predicted macrosomia definition(s)	Outcomes
Aviram 2014	3319	Retro-spective cohort	>37w	Unknown	HC>90th centile (>347mm)	Shoulder dystocia, caesarean section
Bailis 2010	159	Retro-spective cohort	>30w	Back-ground rate	AC>90th centile, HC<50th centile, HC/AC<0.95	Shoulder dystocia
Basit 2010	104	Retro-spective cohort	Unknown	Unknown	EFW>95th centile	Shoulder dystocia, caesarean section, PPH
Bochner 1987	201	Retro-spective cohort	30-33w	All GDM	AC>90th centile	Shoulder dystocia, caesarean section, Birth trauma
Buikema 2014	247	Retro-spective cohort	37-42w	Unknown	EFW>90th centile (all AC>35cm)	Shoulder dystocia
Chaabane 2013	930	Retro-spective case control	37-42w within 72hrs of delivery	Unknown	AC≥350mm	Shoulder dystocia
Chauhan 2011	61	Retro-spective case control	most recent USS after 30w (>90% >36w)	Back-ground rate	EFW≥4500g, EFW≥4000 and <4500, EFW≥90th centile, AC≥35cm	Shoulder dystocia, caesarean section

Chavan 2013	296	Retro-spective cohort	28-36w	All GDM	AC≥75th centile	Shoulder dystocia, neonatal hypoglycaemia, neonatal hyper-bilirubinaemia, 1 min Apgar <7, NICU admission
Cohen 1996	31	Retro-spective cohort	within 2w of delivery	All GDM and pre-existing	AD-BPD≥2.6cm	Shoulder dystocia, brachial plexus injury, neonatal trauma
Cohen 1999	123	Retro-spective cohort	most recent USS after 30w (36w average)	All GDM	Residual AD-BPD (actual-50th centile)=.33-2.0	Shoulder dystocia
Eken 2015	28	Prospective cohort	Unknown	Back-ground Mixed	EFW≥4kg Hadlock, AC>350mm, AC>360mm	Shoulder dystocia, brachial plexus injury
Elliott 1982	70	Retro-spective cohort	within 2w of delivery	All GDM and pre-existing	Chest diameter-BPD≥1.4cm	Shoulder dystocia
Endres 2015	362	Retro-spective matched case-control	within 4w of delivery	Back-ground Matched	AC to HC difference of ≥50mm	Shoulder dystocia, brachial plexus injury
Kehl 2011	258	Prospective cohort	within 3days of delivery >37w	Unknown	AC≥36cm	Caesarean section rate, 2nd/3rd degree tear
Levine 1992	406	Retro-spective cohort	>36w	Back-ground 22% DM	EFW≥90th centile Hadlock	Caesarean section (total), emergency

						caesarean section
Melamed 2010	1938	Prospective czse-control	within 3days of delivery	None	EFW=4000-4999	Shoulder dystocia, caesarean section, 3rd/4th degree tear, PPH, neonatal trauma
Miller 2007	322	Retro-spective cohort	within 2w of delivery	Back-ground 19%	AD minus BPD≥2.6cm, AC>95th centile, EFW>97th centile, EFW>4000g	Shoulder dystocia
Nelson 2009	86	Cohort	28w-32+6w	All GDM and pre-existing	EFW>75th centile	Caesarean section (total), caesarean section for labour arrest
Parry 2000	264	Retro-spective cohort	>37w	Back-ground 6-15%	EFW≥4000g Hadlock	Caesarean section
Peleg 2015	443	Retro-spective cohort	38-42w on admissio n in labour	None	EFW≥4kg Hadlock	Shoulder dystocia, caesarean section (total), emergency caesarean section, 3rd/4th degree tear
Rajan 2009	159	Retro-spective cohort	≥36w	None	AD-BPD≥2.6cm, AC>90th centile	Shoulder dystocia

Scifres 2015	903	Retro-spective cohort	within 31 days of delivery	All GDM	EFW>90th centile Hadlock	Shoulder dystocia, caesarean section, 3rd/4th degree tear, neonatal composite
Simpson 2015	413	Retro-spective cohort	28-32w	All GDM	AC≥70th centile	Caesarean section, neonatal hypoglycaemia
Tatarova 2004	1847	Retro-spective cohort	within 7 days of delivery	Unknown	EFW ≥4000g	Caesarean section
Weiner 2002	315	Prospective cohort	within 3 days of delivery	Back-ground Mixed 9.2% DM	EFW≥4000g Shepard	Shoulder dystocia, caesarean section, Erb's palsy

9.4 Appendix 4: Full text articles reviewed for systematic review

Table 6: List of full text articles retrieved for systematic review: assessment of inclusion/exclusion/reason for exclusion

Author	Title	Year	Published	Type of article	Outcome	Exclusion reason
Abedzadeh-Kalahroudi , M	Incidence of Neonatal Birth Injuries and Related Factors in Kashan, Iran.	2015	Arch Trauma Res	Journal	Exclude	No measure of predicted fetal macrosomia
Abramowicz, J. S	Incorporating sonographic cheek-to-cheek diameter, biparietal diameter and abdominal circumference improves weight estimation in the macrosomic fetus.	1997	Ultrasound in Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Abramowicz, J. S	Fetal cheek-to-cheek diameter in the prediction of mode of delivery.	2005	American Journal of Obstetrics & Gynecology	Journal	Maybe	No response
Acker, D. B.	Risk factors for shoulder dystocia.	1985	Obstetrics & Gynecology	Journal	Exclude	No measure of predicted fetal macrosomia
Assimakopoulos, E.	Fetal abdominal subcutaneous tissue thickness measured by ultrasound at term is associated with birth weight and mode of delivery.	2007	Clin Exp Obstet Gynecol	Journal	Exclude	No measure of predicted fetal macrosomia
Aviram, A	The association between head circumference (HC)> 90th percentile at term and pregnancy outcome.	2014	American Journal of Obstetrics and Gynecology Conference: 34th Annual Meeting of	Conference	Include	

			the Society for Maternal-Fetal Medicine:			
Bailis, A	Accelerated abdominal circumference growth (>75th%ile) in average gestational age fetuses increases risk of shoulder dystocia.	2010	Reproductive Sciences Conference. 57th Annual Scientific Meeting of Society for Gynecologic Investigation, SGI 2010	Conference	Include	
Ballestas, J. D.	Accuracy of a statistical risk assessment model in predicting shoulder dystocia and brachial plexus injury in a high-risk population at Einstein medical center, Philadelphia.	2015	Obstetrics and Gynecology Conference: 63rd Annual Clinical and Scientific Meeting of the American College of Obstetricians and Gynecologists San Francisco	Conference	Maybe	Data not extractable
Bamberg, C.	Fetal macrosomia. Indication for elective cesarean delivery?	2013	Gynakologische Praxis	Journal	Exclude	Lack of original data
Basha, A. S.	Accuracy of sonographic fetal weight estimation within 14 days of delivery in a Jordanian population using Hadlock formula 1.	2012	Medical Principles & Practice	Journal	Exclude	Inappropriate outcome measure
Basit, I	The antenatal and peripartum management of pregnancies with macrosomic babies weighing >5000 g at two tertiary hospitals: A Dublin experience	2010	Archives of Disease in Childhood: Fetal and Neonatal Edition Conference: 14th Annual Conference of the British Maternal and Fetal Medicine Society	Conference	Include	

Bas-Lando, M	Induction at term for women with gestational diabetes mellitus (GDM): Safe for the mother but questionable for the neonate	2012	American Journal of Obstetrics and Gynecology Conference: 32nd Annual Meeting of the Society for Maternal-Fetal Medicine	Conference	Exclude	No measure of predicted fetal macrosomia
Ben-Haroush, A	Accuracy of sonographic estimation of fetal weight before induction of labor in diabetic pregnancies and pregnancies with suspected fetal macrosomia.	2003	Journal of Perinatal Medicine	Journal	Exclude	Inappropriate outcome measure
Ben-Haroush, A	Accuracy of sonographically estimated fetal weight in 840 women with different pregnancy complications prior to induction of labor	2004	Ultrasound in Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Ben-Haroush, A	Fetal weight estimation in diabetic pregnancies and suspected fetal macrosomia	2004	Journal of Perinatal Medicine	Journal	Exclude	Lack of original data
Ben-Haroush, A	Accuracy of a single fetal weight estimation at 29-34 weeks in diabetic pregnancies: can it predict large-for-gestational-age infants at term?	2007	Journal of Perinatal Medicine	Journal	Exclude	Inappropriate outcome measure
Benjamin, S. J	Anthropometric measurements as predictors of cephalopelvic disproportion: Can the diagnostic accuracy be improved?	2012	Acta Obstetricia et Gynecologica Scandinavica	Journal	Exclude	No measure of predicted fetal macrosomia

Blackwell, S	Overestimation of fetal weight by ultrasound: does it influence the likelihood of cesarean delivery for labor arrest?	2009	American Journal of Obstetrics & Gynecology	Journal	Exclude	No measure of predicted fetal macrosomia
Blair, E	Intrauterine growth and spastic cerebral palsy II. The association with morphology at birth.	1992	Early Human Development	Journal	Exclude	No measure of predicted fetal macrosomia
Bo, S	Obesity or diabetes: what is worse for the mother and for the baby?	2003	Diabetes & Metabolism	Journal	Exclude	No measure of predicted fetal macrosomia
Bochner, C. J	Early third-trimester ultrasound screening in gestational diabetes to determine the risk of macrosomia and labor dystocia at term.	1987	American Journal of Obstetrics and Gynecology	Journal	Include	
Bollepalli, S	Asymmetric large-for-gestational-age infants of type 1 diabetic women: morbidity and abdominal growth.	2010	American Journal of Perinatology	Journal	Exclude	Inappropriate outcome measure
Bonnett, T. J	Management of large-for-gestational-age pregnancy and outcomes in non-diabetic fetal macrosomia.	2011	Archives of Disease in Childhood	Conference	Exclude	Inappropriate outcome measure
Boulvain, M	Induction of labour versus expectant management for large-for-date fetuses: a randomised controlled trial.	2015	Lancet	Journal	Maybe	Data not extractable

Boulvain, M.	Induction of labor or expectant management for large-for-dates fetuses: A randomized controlled trial.	2012	American Journal of Obstetrics and Gynecology	Journal	Exclude	Duplicate publication
Bryant, D. R	Limited usefulness of fetal weight in predicting neonatal brachial plexus injury.	1998	Am J Obstet Gynecol	Journal	Exclude	No measure of predicted fetal macrosomia
Buikema, T	Fetal abdominal circumference (AC) of 35cm or greater predicts shoulder dystocia in fetuses presumed to have estimated fetal weight (EFW) appropriate for gestational age (AGA)	2014	American Journal of Obstetrics and Gynecology	Conference	Include	
Buikema, T. A	Can fetal abdominal circumference at term predict shoulder dystocia?	2014	Obstetrics and Gynecology	Conference	Exclude	Duplicate publication
Cater, J. I	Correlates of low birth weight	1980	Child Care Health Dev	Journal	Exclude	Incorrect population
Caughey, A. B	Obstetric ultrasound for estimated fetal weight: is the information more harm than benefit?	2012	American Journal of Obstetrics & Gynecology	Journal	Exclude	Lack of original data
Chaabane, K	Antenatal macrosomia prediction using sonographic fetal abdominal circumference in South Tunisia.	2013	The Pan African medical journal	Journal	Include	
Chauhan, S	Detection of shoulder dystocia among newborns with birth weight of 4,500 g or more: Seven years experience	2011	American Journal of Obstetrics and Gynecology Conference	Conference	Maybe	Include

Chauhan, S. P	Antepartum detection of macrosomic fetus: clinical versus sonographic, including soft-tissue measurements	2000	Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Chauhan, S. P	A scoring system for detection of macrosomia and prediction of shoulder dystocia: a disappointment	2006	Journal of Maternal-Fetal & Neonatal Medicine	Journal	Exclude	Data not extractable
Chavan, N	Utility of fetal abdominal circumference (AC) in predicting disease progression and perinatal outcomes in gestational diabetes mellitus (GDM)	2013	American Journal of Obstetrics and Gynecology Conference	Conference	Include	
Chavan, N	Does an interval change in fetal AC predict perinatal outcomes in GDM?	2014	American Journal of Obstetrics and Gynecology Conference	Conference	Exclude	No measure of predicted fetal macrosomia
Chervenak, J. L	Macrosomia in the postdate pregnancy: is routine ultrasonographic screening indicated?	1989	American Journal of Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Cody, F.	Can ultrasound-estimated fetal weight accurately predict birthweight near term in singleton nulliparous patients? Results from the Multicenter Prospective Genesis Study	2016	BJOG: An International Journal of Obstetrics and Gynaecology Conference	Conference	Exclude	Inappropriate outcome measure
Cohen, B.	Sonographic prediction of shoulder dystocia in infants of diabetic mothers.	1996	Obstet Gynecol	Journal	Include	

Cohen, B. F	The incidence and severity of shoulder dystocia correlates with a sonographic measurement of asymmetry in patients with diabetes	1999	Am J Perinatol	Journal	Include	
Combs, C. A	Elective induction versus spontaneous labor after sonographic diagnosis of fetal macrosomia	1993	Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Conway, D. L	Elective delivery of infants with macrosomia in diabetic women: reduced shoulder dystocia versus increased cesarean deliveries	1998	American Journal of Obstetrics & Gynecology	Journal	Maybe	Data not extractable
Coomarasamy, A	Accuracy of ultrasound biometry in the prediction of macrosomia: a systematic quantitative review.	2005	BJOG: An International Journal of Obstetrics & Gynaecology	Journal	Exclude	Lack of original data
Corcoran, S	Accuracy of third trimester ultrasound in predicting macrosomia in patients with diabetes in pregnancy	2014	American Journal of Obstetrics and Gynecology Conference	Conference	Exclude	Inappropriate outcome measure
Culligan, P. J	Elective cesarean section to prevent anal incontinence and brachial plexus injuries associated with macrosomia--a decision analysis	2005	International Urogynecology Journal	Journal	Exclude	Lack of original data
Culliney, K	Regimens of fetal surveillance of suspected large-for-gestational-age fetuses for improving health outcomes	2016	Cochrane Database of Systematic Reviews	Journal	Exclude	Lack of original data

Dadkhah, F	Predicting neonatal weight of more than 4000 g using fetal abdominal circumference measurement by ultrasound at 38-40 weeks of pregnancy: a study in Iran	2013	Journal of Obstetrics & Gynaecology Research	Journal	Exclude	Inappropriate outcome measure
Day, S. M	Birth size and the risk of cerebral palsy in term births	2014	Obstetrics & Gynecology	Journal	Exclude	Lack of original data
Delpapa, E. H	Pregnancy outcome following ultrasound diagnosis of macrosomia	1991	Obstetrics & Gynecology	Journal	Maybe	Data not extractable
Donnelley, E. L.	Body composition of large for gestational age neonates measured using the Peapod	2012	Journal of Paediatrics and Child Health	Conference	Exclude	No measure of predicted fetal macrosomia
Doubilet, P. M	Fetal growth disturbances	1990	Seminars in Roentgenology	Journal	Exclude	Lack of original data
Drossou, V	Accuracy of anthropometric measurements in predicting symptomatic SGA and LGA neonates	1995	Acta Paediatr	Journal	Exclude	No measure of predicted fetal macrosomia
Duryea, E. L	The utility of the femur length to abdominal circumference ratio for the prediction of shoulder dystocia and neonatal brachial plexus palsy in women with gestational diabetes	2016	American Journal of Obstetrics and Gynecology	Conference	Maybe	Data not extractable
Dyachenko, A	Prediction of risk for shoulder dystocia with neonatal injury	2006	Am J Obstet Gynecol	Journal	Exclude	No measure of predicted fetal macrosomia

Eik-Nes, S. H	Routine ultrasound fetal examination in pregnancy: the 'Alesund' randomized controlled trial	2000	Ultrasound in obstetrics & gynecology	Journal	Exclude	No measure of predicted fetal macrosomia
Eken, M	Six-year incidence and some features of cases of brachial plexus injury in a tertiary referral center	2015	Turk Jinekoloji ve Obstetrik Dernegi Dergisi	Journal	Include	
Elliott, J. P	Ultrasonic prediction of fetal macrosomia in diabetic patients	1982	Obstet Gynecol	Journal	Include	
Endres, L.	Association of fetal abdominal-head circumference size difference with shoulder dystocia: A multicenter study	2015	AJP Reports	Journal	Include	
Erhardt, C. L	Influence of weight and gestation on perinatal and neonatal mortality by ethnic group	1964	Am J Public Health Nations Health	Journal	Exclude	Incorrect population
Eze, C. U	Correlation of ultrasonographic estimated fetal weight with actual birth weight in a tertiary hospital in Lagos, Nigeria	2015	African Health Sciences	Journal	Exclude	Inappropriate outcome measure
Eziefule, A. A	Diabetic delivery of newborns with birthweight > 4,500 grams: Detection may be problematic	2016	American Journal of Obstetrics and Gynecology	Conference	Exclude	Inappropriate outcome measure
Faschingbauer, F	Intrapartum sonographic weight estimation	2015	Archives of Gynecology & Obstetrics	Journal	Exclude	Inappropriate outcome measure
Foroutan, J	Does the timing of ultrasound affect the ability to predict birthweight in gestational diabetes?	2014	Reproductive Sciences	Conference	Exclude	Inappropriate outcome measure

Fournie, A	Induce or not induce labor in gestational diabetes	2002	Journal de Gynecologie Obstetrique et Biologie de la Reproduction	Journal	Translate	Lack of original data
Froehlich, R. J	The association of estimated fetal weight and cesarean delivery in women attempting vaginal delivery at term	2016	American Journal of Obstetrics and Gynecology	Conference	Maybe	Data not extractable
Garabedian, C	Detection of fetal macrosomia in diabetic pregnancies: Evaluation of the measurements of subcutaneous tissue width and fetal liver by a longitudinal follow-up during the 2nd and the 3rd trimester	2012	Journal of Maternal-Fetal and Neonatal Medicine	Conference	Exclude	Inappropriate outcome measure
Geifman-Holtzman, O	The clinical utility of oral glucose tolerance test at term: can it predict fetal macrosomia?	2010	Archives of Gynecology & Obstetrics	Journal	Exclude	Data not extractable
Gherman, R	Fetal abdominal circumference and macrosomia	2001	Journal of Reproductive Medicine	Journal	Exclude	Lack of original data
Gilby, J. R	Fetal abdominal circumference measurements of 35 and 38 cm as predictors of macrosomia. A risk factor for shoulder dystocia	2000	Journal of Reproductive Medicine	Journal	Exclude	Inappropriate outcome measure
Goh, J. E	Metformin for gestational diabetes in routine clinical practice	2011	Diabetic Medicine	Journal	Exclude	No measure of predicted fetal macrosomia

Gonen, R	Is macrosomia predictable, and are shoulder dystocia and birth trauma preventable?	1996	Obstetrics and Gynecology	Journal	Maybe	Data not extractable
Hackmon, R	Combined analysis with amniotic fluid index and estimated fetal weight for prediction of severe macrosomia at birth	2007	American Journal of Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Haram, K	Suspected large fetus in the last period of pregnancy--a difficult problem	2001	Tidsskrift for Den Norske Laegeforening	Journal	Exclude	Lack of original data
Haram, K	Suspected big baby: a difficult clinical problem in obstetrics	2001	Acta Obstetricia et Gynecologica Scandinavica	Journal	Exclude	Duplicate publication
Hawarden, A	Gestational diabetes in Chester: Comparison throughout the years	2016	BJOG: An International Journal of Obstetrics and Gynaecology	Conference	Exclude	Inappropriate outcome measure
Hehir, M. P	Increased fetal adiposity is a risk factor for cesarean delivery - Results of the national prospective Genesis Study	2016	American Journal of Obstetrics and Gynecology	Journal	Maybe	No data provided by author
Henrichs, C	Detecting fetal macrosomia with abdominal circumference alone	2003	Journal of Reproductive Medicine	Journal	Exclude	Inappropriate outcome measure
Heo, H. J	Sonographic suspicion of macrosomia effecting mode of delivery	2009	American Journal of Obstetrics and Gynecology	Conference	Maybe	Data not extractable

Hernandez-Martinez, A	Predictive model for risk of cesarean section in pregnant women after induction of labor	2016	Arch Gynecol Obstet	Journal	Exclude	No measure of predicted fetal macrosomia
Hitschold, T	Shoulder dystocia--risk factors and indicators. A retrospective analysis within an obstetric collective of 14,913 deliveries	2008	Z Geburtshilfe Neonatol	Journal	Exclude	No measure of predicted fetal macrosomia
Hod, M	Perinatal complications following gestational diabetes mellitus' how 'sweet' is ill?	1996	Acta Obstetricia et Gynecologica Scandinavica	Journal	Exclude	No measure of predicted fetal macrosomia
Hod, M	Antepartum management protocol. Timing and mode of delivery in gestational diabetes	1998	Diabetes Care	Journal	Exclude	Lack of original data
Hoffert Gilmartin, A. B	Gestational diabetes mellitus	2008	Reviews in Obstetrics & Gynaecology	Journal	Exclude	Lack of original data
Holcomb, W. L	Abdominal circumference vs. estimated weight to predict large for gestational age birth weight in diabetic pregnancy	2000	Clinical Imaging	Journal	Exclude	Inappropriate outcome measure
Holubiev, A. P	Comparative Evaluation of Some Methods of Determining the Weight of the Intrauterine Fetus	1963	Pediatriia, Akusherstvo i Ginekologiia	Journal	Exclude	No measure of predicted fetal macrosomia
Hsu, T. Y	Neonatal clavicular fracture: clinical analysis of incidence, predisposing factors, diagnosis, and outcome	2002	Am J Perinatol	Journal	Exclude	No measure of predicted fetal macrosomia

Iffy, L	The risk of shoulder dystocia related permanent fetal injury in relation to birth weight	2008	Eur J Obstet Gynecol Reprod Biol	Journal	Exclude	No measure of predicted fetal macrosomia
Iftikhar, R	Intrapartum complications of macrosomic fetus	2007	Journal of the Liaquat University of Medical and Health Sciences	Journal	Exclude	No measure of predicted fetal macrosomia
Ismail, M.	Birthweight prediction by clinical versus ultrasound examination estimation among pregnant mothers in hospital Universiti Sains Malaysia, Kelantan Malaysia	2013	BJOG: An International Journal of Obstetrics and Gynaecology	Conference	Maybe	Data not extractable
Jaffe, R	Identification of fetal growth abnormalities in diabetes mellitus	2002	Semin Perinatol	Journal	Exclude	Lack of original data
Jango, H	Modifiable risk factors of obstetric anal sphincter injury in primiparous women: A population-based cohort study	2014	American Journal of Obstetrics and Gynecology	Journal	Exclude	No measure of predicted fetal macrosomia
Jazayeri, A	Macrosomia prediction using ultrasound fetal abdominal circumference of 35 centimeters or more	1999	Obstet Gynecol	Journal	Exclude	Data not extractable
Johnstone, F. D	Clinical and ultrasound prediction of macrosomia in diabetic pregnancy.	1997	International Journal of Gynecology & Obstetrics	Journal	Exclude	Duplicate publication
Johnstone, F. D	Prediction of macrosomia in diabetic pregnancy	1997	Contemporary Reviews in Obstetrics and Gynaecology	Journal	Exclude	Inappropriate outcome measure

Kamanu, C. I	Fetal macrosomia in African women: a study of 249 cases	2009	Arch Gynecol Obstet	Journal	Exclude	No measure of predicted fetal macrosomia
Kase, B	Can the addition of customized standards improve the prediction of fetal macrosomia and cesarean delivery for arrested labor?	2012	American Journal of Obstetrics and Gynecology	Conference	Maybe	Data not extractable
Kees, S	Features of shoulder dystocia in a busy obstetric unit	2001	J Reprod Med	Journal	Maybe	Data not extractable
Kehl, S	Role of fetal abdominal circumference as a prognostic parameter of perinatal complications	2011	Arch Gynecol Obstet	Journal	Include	
Keller, J. D	Infants of diabetic mothers with accelerated fetal growth by ultrasonography: are they all alike?	1990	American Journal of Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Kernaghan, D	Fetal size and growth velocity in the prediction of the large for gestational age (LGA) infant in a glucose impaired population	2007	European Journal of Obstetrics, Gynecology, & Reproductive Biology	Journal	Maybe	No response from author
King, J	Increased maternal and neonatal morbidity associated with suspected fetal macrosomia	2012	American Journal of Obstetrics and Gynecology	Journal	Maybe	Data not extractable
King, J. R	Increased composite maternal and neonatal morbidity associated with ultrasonographically suspected fetal macrosomia.	2012	J Matern Fetal Neonatal Med	Journal	Exclude	Duplicate publication

Kitzmiller, J. L	Randomized trial evaluating a predominately fetal growth-based strategy to guide management of gestational diabetes in Caucasian women: Response to Schaefer-Graf et al	2004	Diabetes Care	Journal	Exclude	Lack of original data
Kjos, S. L	Modified therapy for gestational diabetes using high-risk and low-risk fetal abdominal circumference growth to select strict versus relaxed maternal glycemic targets	2007	Diabetes Care	Journal	Exclude	Lack of original data
Klaij, F. A	Humerospinous distance measurements: accuracy and usefulness for predicting shoulder dystocia in delivery at term	1998	Ultrasound Obstet Gynecol	Journal	Exclude	No measure of predicted fetal macrosomia
Kongnyuy, E. J	Comment on "Predicting term birth weight using ultrasound and maternal characteristics	2007	Eur J Obstet Gynecol Reprod Biol	Journal	Exclude	Lack of original data
Kranz, C	Impact and limitations of ultrasound examination immediately before delivery--a prospective evaluation with 1127 pregnant women	2011	Z Geburtshilfe Neonatol	Journal	Translate	Inappropriate outcome measure
Lao, T. T	Perinatal outcome in large-for-gestational-age infants. Is it influenced by gestational impaired glucose tolerance?	2002	J Reprod Med	Journal	Exclude	No measure of predicted fetal macrosomia

Larkin, J	Use of a customized growth standard to identify LGA neonates predicts intrapartum morbidity.	2010	Reproductive Sciences	Conference	Exclude	Duplicate publication
Larkin, J	An alternative method for customization of fetal growth standards	2011	American Journal of Obstetrics and Gynecology	Conference	Exclude	Duplicate publication
Larkin, J	An individualized fetal growth trajectory improves identification of macrosomia-related risk.	2011	American Journal of Obstetrics and Gynecology	Conference	Maybe	No response from author
Larson, A	Head circumference at birth and risk for shoulder dystocia during delivery of macrosomic infants	2010	Annals of Neurology	Conference	Exclude	No measure of predicted fetal macrosomia
Lee, B. H	Association between fetal abdominal circumference and birthweight in maternal hyperglycemia	2014	Acta Obstetricia et Gynecologica Scandinavica	Journal	Exclude	Inappropriate outcome measure
Lepercq, J	Factors associated with cesarean delivery in nulliparous women with type 1 diabetes	2010	Obstet Gynecol	Journal	Exclude	Incorrect population
Levine, A. B	Sonographic diagnosis of the large for gestational age fetus at term: does it make a difference?	1992	Obstetrics & Gynecology	Journal	Include	
Li, N	Risk factors and clinical prediction of shoulder dystocia in non-macrosomia	2015	Zhonghua Fu Chan Ke Za Zhi	Journal	Exclude	No measure of predicted fetal macrosomia
Liang, J. Z	Developing parameters for predicting macrosomia	2008	Sichuan da Xue Bao Medical Science Medical Science		Exclude	Inappropriate outcome measure

Liberty, G	A proportion score of pelvic and neonatal head circumference is highly predictive of instrumental delivery and cesarean section due to cephalo-pelvic disproportion	2013	American Journal of Obstetrics and Gynecology	Conference	Exclude	No measure of predicted fetal macrosomia
Little, S	Does obtaining an ultrasound estimated fetal weight near term increase the risk of cesarean delivery?	2012	American Journal of Obstetrics and Gynecology	Conference	Maybe	Data not extractable
Little, S. E	Estimated fetal weight by ultrasound: a modifiable risk factor for cesarean delivery?	2012	Am J Obstet Gynecol	Journal	Exclude	Duplicate publication
Lukacin, S	Birth weight. body length, and head circumference of the newborn from the 32nd to the 43rd week of pregnancy	1976	Bratisl Lek Listy	Journal	Exclude	No measure of predicted fetal macrosomia
Lunell, N. O	Potential diabetes in women with large babies. A follow-up study	1972	Acta Obstet Gynecol Scand	Journal	Exclude	No measure of predicted fetal macrosomia
Malfertheiner, S. F	Gestational diabetes, comparison of the management and neonatal outcome over the time of a decade	2010	Archives of Gynecology and Obstetrics	Conference	Exclude	No measure of predicted fetal macrosomia
Malin, G. L	Comparison of mri and ultrasound to detect fetal macrosomia at term: A systematic review and meta-analysis	2014	Archives of Disease in Childhood: Fetal and Neonatal Edition	Journal	Exclude	Lack of original data

Malin, G. L	Antenatal magnetic resonance imaging versus ultrasound for predicting neonatal macrosomia: a systematic review and meta-analysis	2016	BJOG: An International Journal of Obstetrics & Gynaecology	Journal	Exclude	Lack of original data
Malinowska-Polubiec, A	Shoulder dystocia in diabetic and non-diabetic pregnancies	2014	Neuro Endocrinol Lett	Journal	Exclude	No measure of predicted fetal macrosomia
Marsh, M	BJOG editor's choice	2013	BJOG: An International Journal of Obstetrics and Gynaecology	Journal	Exclude	Lack of original data
Mattoo, G. M	Maximum thigh circumference as an indicator of birth weight	1991	Indian J Matern Child Health	Journal	Exclude	No measure of predicted fetal macrosomia
Mazouni, C	Development and internal validation of a nomogram to predict macrosomia	2007	Ultrasound in Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
McEwan, H. P	The oversized baby. A study of 169 cases	1966	Journal of Obstetrics & Gynaecology of the British Commonwealth	Journal	Exclude	No measure of predicted fetal macrosomia
Mehta, S. H	Accuracy of estimated fetal weight in shoulder dystocia and neonatal birth injury	2005	Am J Obstet Gynecol	Journal	Exclude	No measure of predicted fetal macrosomia
Melamed, N.	Sonographic prediction of fetal macrosomia: the consequences of false diagnosis	2010	J Ultrasound Med	Journal	Include	

Melendez, J	Severe shoulder dystocia leading to neonatal injury: a case control study	2009	Arch Gynecol Obstet	Journal	Exclude	No measure of predicted fetal macrosomia
Miller, J. M	Fetal overgrowth. Diabetic versus nondiabetic	1988	Journal of Ultrasound in Medicine	Journal	Exclude	Inappropriate outcome measure
Miller, R. S	Sonographic fetal asymmetry predicts shoulder dystocia	2007	Journal of Ultrasound in Medicine	Journal	Include	
Mimouni, F	Birth trauma in insulin-dependent diabetic pregnancies	1992	Am J Perinatol	Journal	Exclude	No measure of predicted fetal macrosomia
Mirghani, H	Fetal body measurement in gestational diabetes	2009	International Journal of Gynecology and Obstetrics	Journal	Exclude	No measure of predicted fetal macrosomia
Morales-Rosello, J	Study of abdominal circumference proportions in fetuses with growth disorders	2005	Archives of Gynecology & Obstetrics	Journal	Exclude	Inappropriate outcome measure
Morgan, M. A.	Efficacy of the fetal-pelvic index in nulliparous women at high risk for fetal-pelvic disproportion	1992	American Journal of Obstetrics & Gynecology	Journal	Exclude	Incorrect population
Mtimavalye, L. A	The relationship between increasing birthweight and cephalopelvic disproportion in Dar Es Salaam, Tanzania	1974	Journal of Obstetrics & Gynaecology of the British Commonwealth	Journal	Exclude	No measure of predicted fetal macrosomia
Navti, O. B	The peri-partum management of pregnancies with macrosomic babies weighing > or =4,500 g at a tertiary University Hospital	2007	J Obstet Gynaecol	Journal	Exclude	No measure of predicted fetal macrosomia

Nelson, L	Prediction of fetal macrosomia in term infants of diabetic mothers based on early third trimester ultrasound	2009	American Journal of Obstetrics and Gynecology	Conference	Include	
Nelson, L	Prediction of large for gestational age birth weights in diabetic mothers based on early third-trimester sonography	2011	Journal of Ultrasound in Medicine	Journal	Exclude	Duplicate publication
Njoku, C	Determination of accuracy of fetal weight using ultrasound and clinical fetal weight estimations in calabar, Cross River State, Nigeria	2015	International Journal of Gynecology and Obstetrics Conference	Conference	Exclude	Inappropriate outcome measure
Okun, N	Relative importance of maternal constitutional factors and glucose intolerance of pregnancy in the development of newborn macrosomia	1997	Journal of Maternal-Fetal Medicine	Journal	Exclude	No measure of predicted fetal macrosomia
O'Reilly-Green, C. P	Receiver operating characteristic curves of sonographic estimated fetal weight for prediction of macrosomia in prolonged pregnancies	1997	Ultrasound in Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Ouzounian, J	Clinical risk factors do not predict shoulder dystocia	2014	American Journal of Obstetrics and Gynecology	Conference	Exclude	No measure of predicted fetal macrosomia
Palatnik, A	Predictors of shoulder dystocia at the time of operative vaginal delivery	2016	American Journal of Obstetrics and Gynecology	Conference	Exclude	No measure of predicted fetal macrosomia

Parra Saavedra, M. A	Detection of late pregnancy Large-for-gestational age fetuses (LGA) by ultrasound and Doppler evaluation at routine third trimester scan in adequate-for-gestational age (AGA) pregnancies	2015	Journal of Perinatal Medicine	Conference	Exclude	No measure of predicted fetal macrosomia
Parry, S	Ultrasonographic prediction of fetal macrosomia. Association with cesarean delivery	2000	J Reprod Med	Journal	Include	
Pavicic, M	Do HAPO based glycemic targets change the ability of an enlarged fetal abdominal circumference (AC) at 28-34 weeks gestation to predict obstetric and neonatal morbidity in gestational diabetes (GDM)?	2015	International Journal of Gynecology and Obstetrics	Journal	Maybe	No data received from author
Peleg, D	The effect of fetal weight estimation on cesarean section rates of the macrosomic neonate	2014	American Journal of Obstetrics and Gynecology	Conference	Exclude	Duplicate publication
Peleg, D	Counseling for fetal macrosomia: an estimated fetal weight of 4,000 g is excessively low	2015	Am J Perinatol	Journal	Include	
Pregazzi, R	Prevention of obstetrical damage to the perineum. Proposal for a fetal-pelvic index for the selection of parturients at risk	1998	Minerva Ginecologica	Journal	Translate	No measure of predicted fetal macrosomia
Quevedo, S. F	Treatment of gestational diabetes complicated by significant maternal	2010	Diabetes	Conference	Exclude	No measure of predicted fetal macrosomia

	obesity: Utility of fetal-based therapy					
Rajan, P. V.	Correlation of increased fetal asymmetry with shoulder dystocia in the nondiabetic woman with suspected macrosomia	2009	J Reprod Med	Journal	Include	
Raychaudhuri, K	Glycemic control throughout pregnancy and fetal growth in insulin-dependent diabetes	2000	Obstetrics and Gynecology	Journal	Exclude	No measure of predicted fetal macrosomia
Redfearn, C	Accuracy of ultrasound-estimated fetal weight in large for gestational age/macrosomic infants	2016	BJOG: An International Journal of Obstetrics and Gynaecology	Conference	Exclude	Inappropriate outcome measure
Redfearn, C	Delivery outcomes of ultrasound scans indicated large for gestational age infants (>4000 g)	2016	BJOG: An International Journal of Obstetrics and Gynaecology	Conference	Maybe	No data received from author
Reid E. W	Macrosomia in uncomplicated pregnancies: an investigation of physical activity, maternal nutrition and women's experiences.	2012	Queen's University, Belfast	PhD	Exclude	No measure of predicted fetal macrosomia
Rosati, P	The role of echography in the management of fetal macrosomia	1995	Minerva Ginecol	Journal	Translate	No measure of predicted fetal macrosomia
Rossi, G	Adequate timing of fetal ultrasound to guide metabolic therapy in mild gestational diabetes mellitus. Results from a randomized study	2000	Acta Obstetricia et Gynecologica Scandinavica	Journal	Maybe	Data not extractable

Rouse, D. J	The effectiveness and costs of elective cesarean delivery for fetal macrosomia diagnosed by ultrasound	1996	JAMA	Journal	Exclude	Lack of original data
Rouse, D. J	Prophylactic cesarean delivery for fetal macrosomia diagnosed by means of ultrasonography--A Faustian bargain?	1999	Am J Obstet Gynecol	Journal	Exclude	Lack of original data
Rouse, D. J.	Sonography, suspected macrosomia, and prophylactic cesarean: a limited partnership	2000	Clin Obstet Gynecol	Journal	Exclude	Lack of original data
Sanchez-Ramos, L	Expectant management versus labor induction for suspected fetal macrosomia: a systematic review	2002	Obstetrics and Gynecology	Journal	Exclude	Lack of original data
Sandmire, H. F	Whither ultrasonic prediction of fetal macrosomia?	1993	Obstetrics & Gynecology	Journal	Exclude	Lack of original data
Sator-Kulic, Z	Correlation between certain parameters in pregnancy and body weight and the vitality of the neonate	1990	Med Pregl	Journal	Exclude	No measure of predicted fetal macrosomia
Schwartz, R	Elective delivery of large infants in diabetic pregnancies: Consequences and implications	2011	Reproductive Sciences	Conference	Exclude	No measure of predicted fetal macrosomia
Schwartz, R. A	Can abdominal circumference in the second trimester predict the large for gestational age neonate?	2012	Reproductive Sciences	Conference	Exclude	Inappropriate outcome measure

Scifres, C	The impact of a large for gestational age diagnosis on the risk for cesarean delivery in women with gestational diabetes	2015	American Journal of Obstetrics and Gynecology	Conference	Maybe	Data not extractable
Scifres, C. M	Large-for-Gestational-Age Ultrasound Diagnosis and Risk for Cesarean Delivery in Women With Gestational Diabetes Mellitus	2015	Obstet Gynecol	Journal	Include	
Sentilhes, L	Shoulder dystocia: guidelines for clinical practice from the French College of Gynecologists and Obstetricians (CNGOF)	2016	Eur J Obstet Gynecol Reprod Biol	Journal	Exclude	Lack of original data
Seyam, Y. S	Ultrasound prediction of fetal macrosomia in diabetic women and its effect on the route of delivery and the outcome of pregnancy	2007	Qatar Medical Journal	Journal	Exclude	Incorrect population
Shami, S. A	Consanguinity, gestational period and anthropometric measurements at birth in Pakistan	1991	Ann Hum Biol	Journal	Exclude	Incorrect population
Sharifi, F	Association of maternal body mass index (BMI) with adverse prenatal outcomes	2013	Iranian Journal of Reproductive Medicine	Conference	Exclude	Lack of original data
Shi, C. Y	Relationship between fetal abdominal circumference and birth weight: clinical findings in 1475 pregnancies	2005	Zhonghua Fu Chan Ke Za Zhi	Journal	Exclude	Paper unavailable

Simmonds, L	Can ultrasound predict increased neonatal body fat in gestational diabetes?	2010	Reproductive Sciences	Conference	Exclude	Inappropriate outcome measure
Simpson, K	Does an enlarged abdominal circumference predict morbidity for patients diagnosed with gestational diabetes (GDM) using the IADPSG criteria?	2015	International Journal of Gynecology and Obstetrics	Conference	Maybe	Include
Sirota, I	Shoulder dystocia risk factor analysis in non-diabetic versus diabetic pregnant patients: Is induction of labor the right choice?	2013	Reproductive Sciences	Conference	Maybe	Data not extractable
Sood, A. K	Prediction of fetal macrosomia using humeral soft tissue thickness	1995	Obstet Gynecol	Journal	Exclude	Inappropriate outcome measure
Sovio, U	Accelerated fetal growth prior to diagnosis of gestational diabetes mellitus: A prospective cohort study of nulliparous women	2016	Diabetes Care	Journal	Exclude	Inappropriate outcome measure
Stanislaw, H	Comment on "Predicting term birth weight using ultrasound and maternal characteristics	2007	Journal of Reproductive Medicine	Journal	Exclude	Lack of original data
Stanislaw, H.	Accuracy of birth weight prediction methods	2008	American Journal of Obstetrics and Gynecology	Journal	Exclude	Lack of original data
Stein, W	Prediction of shoulder dystocia combining foetal weight estimation	2009	Zeitschrift fur Geburtshilfe und Neonatologie	Conference	Translate	No measure of fetal macrosomia

	by ultrasound and maternal risk factors A solution for the dilemma?					
Tadesse, W. G	Predictors of babies weighing more than 4 kg in pregnancies complicated by diabetes	2016	Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Tamura, R. K	Diabetic macrosomia: accuracy of third trimester ultrasound	1986	Akusherstvo i Ginekologiia	Journal	Exclude	Inappropriate outcome measure
Tatarova, S	Fetal macrosomia: mode of delivery	2004	Diabetes	Conference	Translate	Include
Valent, A	Accuracy of the estimation of fetal weight near delivery by ultrasound in diabetic and nondiabetic pregnancies	2015	Diabetic Medicine	Journal	Exclude	Inappropriate outcome measure
Venkataraman, H	Early ethnic differences in fetal growth patterns in gestational diabetes (GDM)	2015	Diabetes UK Professional	Conference	Exclude	No measure of predicted fetal macrosomia
Vladareanu, S	Suspected fetal macrosomia and the risk of shoulder dystocia as an indication for cesarean section	2015	Gineco.eu	Journal	Exclude	Lack of original data
Voormolen D, N	President's Plenary Session	2014	Reproductive Sciences	Conference	Exclude	No measure of predicted fetal macrosomia
Walsh, J. M	Prediction and prevention of the macrosomic fetus	2012	European Journal of Obstetrics, Gynecology, & Reproductive Biology	Journal	Exclude	Lack of original data
Wang, X	Multicenter analysis of risk factors and clinical characteristics of shoulder dystocia	2015	Zhonghua Fu Chan Ke Za Zhi	Journal	Exclude	No measure of predicted fetal macrosomia

Weeks, J. W	Fetal macrosomia: does antenatal prediction affect delivery route and birth outcome?	1995	Am J Obstet Gynecol	Journal	Exclude	Data not extractable
Weiner, E	Comparison of three methods for estimated fetal weight during active labor: A prospective study	2015	American Journal of Obstetrics and Gynecology	Conference	Exclude	Inappropriate outcome measure
Weiner, E	Comparison between three methods for the detection of macrosomia and growth restriction in patients presenting in active labor-a prospective study	2016	American Journal of Obstetrics and Gynecology	Conference	Exclude	Inappropriate outcome measure
Weiner, Z	Clinical and ultrasonographic weight estimation in large for gestational age fetus	2002	Eur J Obstet Gynecol Reprod Biol	Journal	Include	
Weissmann-Brenner, A	Does the availability of maternal HbA1c results improve the accuracy of sonographic diagnosis of macrosomia?	2004	Ultrasound in Obstetrics & Gynecology	Journal	Exclude	Inappropriate outcome measure
Wennerholm, U. B	Induction of labor versus expectant management for post-date pregnancy: is there sufficient evidence for a change in clinical practice?	2009	Acta Obstetricia et Gynecologica Scandinavica	Journal	Exclude	Lack of original data
Witkop, C. T	Active compared with expectant delivery management in women with gestational diabetes: a systematic review	2009	Obstet Gynecol	Journal	Exclude	Lack of original data

Wiznitzer, A	Obstructed labor and shoulder dystocia	1995	Curr Opin Obstet Gynecol	Journal	Exclude	Lack of original data
Wollschlaeger, K	A study of fetal macrosomia	1999	Arch Gynecol Obstet	Journal	Exclude	No measure of predicted fetal macrosomia
Wright, L. A	1, 5-Anhydroglucitol (1, 5 AG) and neonatal complications in pregnancy complicated by diabetes	2013	Diabetologia	Conference	Exclude	No measure of predicted fetal macrosomia
Yang, J. W	Predictors of fetal macrosomia and cesarean c-section in women with gestational diabetes mellitus	2013	Endocrine Reviews	Conference	Exclude	No measure of predicted fetal macrosomia
Yeo, G. S	An analysis of risk factors for the prediction of shoulder dystocia in 16,471 consecutive births	1995	Ann Acad Med Singapore	Journal	Exclude	No measure of predicted fetal macrosomia

9.5 Appendix 5: Publications from this research

Publications

Rebecca Robinson, Kate F. Walker, Victoria A. White, George J. Bugg, Kym I.E. Snell, Nia W.Jones. **The test accuracy of antenatal ultrasound definitions of fetal macrosomia to predict birth injury: A systematic review.** European Journal of Obstetrics & Gynaecology and Reproductive Biology, March 2020:246;79-85.

Posters and published abstracts

Robinson R, White, V, Bugg, G, Jones, N. **The test accuracy of antenatal ultrasound definitions of fetal macrosomia to predict birth injury. A systematic review and meta-analysis**, BMFMS 2018, Brighton, 19th–20th April 2018. Abstract and Poster.

Robinson R, White V, Bugg G, Jones N. **Big Babies: How big is too big?** Learning in the Library: Student research Showcase. University of Nottingham libraries research poster exhibition, Feb-Mar 2018. Poster.

Robinson R, White, V, Bugg, G, Jones, N. **The test accuracy of antenatal ultrasound definitions of fetal macrosomia to predict birth injury. A systematic review and meta-analysis**, Blair Bell Annual Academic Meeting, RCOG, 8th–9th February 2018. Poster.

Robinson R, White V, Bugg G, Jones N. **Big Babies: How big is too big?** University of Nottingham, School of Medicine Research Showcase, 28th June 2017. Poster and Presentation.

Robinson R, White V, Bugg G, Jones N. **Big Babies: How big is too big?** University of Nottingham Research Showcase, 21st June 2017.
Poster and Presentation.

Prizes

June 2017 1st Prize Poster Presentation: University of Nottingham, School of Medicine Research Showcase

June 2017 Popular Choice Award: University of Nottingham Research Showcase

June 2017 1st Prize: University of Nottingham Research Showcase.