ULTRASOUND DIAGNOSIS OF MACROSOMIA AMONG WOMEN WITH GESTATIONAL DIABETES – REVIEW OF THE LITERATURE

Boris Lovrić1, Siniša Šijanović2, Ivan Zmijanović3, Gordana Jurić4 and Josip Juras5

1Department of Obstetrics and Gynecology, Nova Gradiška General Hospital, Nova Gradiška, Croatia; 2Department of Obstetrics and Gynecology, Osijek University Hospital Center, Faculty of Medicine, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia; 3Department of Obstetrics and Gynecology, Šibenik General Hospital, Šibenik, Croatia; 4Croatian Catholic University, Department of Nursery, Zagreb, Croatia; 5Department of Obstetrics and Gynecology, Zagreb University Hospital Center, School of Medicine, University of Zagreb, Zagreb, Croatia

SUMMARY – Pregnancies burdened with gestational diabetes (GDM) are more likely to end in birth of a macrosomic child, where the frequency of operative termination of pregnancy is more common, accompanied with more complications and injuries of both mother and child in comparison to the general population. The need to calculate fetal weight right before delivery has led to the development of numerous methods for greater estimation accuracy. We reviewed the related literature from 1980 to 2020, using the terms macrosomia, ultrasound assessment, gestational diabetes, and relevant articles were considered in preparation of this article. The most frequently used methods are based on two-dimensional ultrasound measurements of individual fetal biometric parameters and their combination in a mathematical regression model. Some methods involve the addition of other mother and child conditions to increase reliability of the method in recognizing macrosomia. In daily work, especially with pregnant women suffering from GDM, it is necessary to have reliable data on the estimated fetal weight before making the correct clinical decision on how to terminate the pregnancy. In this regard, we bring a review of the literature related to the assessment of fetal macrosomia, especially in women with GDM.

Key words: Gestational diabetes mellitus; Fetal weight; Ultrasound; Assessment; Macrosomia

Introduction

A precisely estimated fetal weight within few days prior to delivery can contribute to clinical decision on how to complete the pregnancy. Previous research related to comparison of the accuracy of different models of fetal weight assessment has not provided an unambiguous answer. We reviewed the literature relevant to the assessment of fetal weight in the group of women with gestational diabetes (GDM), given the higher frequency of macrosomia accompanied with GDM. An increased number of complications arise during pregnancy burdened by GDM, and perinatal outcome is worse in women with GDM, especially if women were obese before pregnancy, as compared with women whose pregnancies were not burdened with GDM1.
Methods

We reviewed relevant literature from 1980 to 2020. The terms used in the literature search in the PubMed database (https://pubmed.ncbi.nlm.nih.gov/) were ‘macrosomia’, ‘ultrasound assessment’ and ‘gestational diabetes’. We compared and described studies that compared diagnostic accuracy of different formulas for fetal weight and/or macrosomia estimation.

Results

Ultrasound assessment of fetal macrosomia

When assessing fetal weight, few seemingly simple questions must be answered, i.e., which method to use, what and when to measure, and what is the probability of clinically relevant error? Today’s standard is two-dimensional (2D) ultrasound estimation. The most widely used method is the Hadlock method, which includes measuring biparietal diameter (BPD), abdominal circumference (AC), and femur length (FL). Many other methods have been developed over time. In the analysis of 18 methods, the authors concluded that most methods were not sufficiently precise when estimating fetal weight over 4000 grams and that new, more precise methods should be developed.

Another study compared 10 fetal weight estimation methods, where the presumed fetal weight was equal to or greater than 4500 grams. The authors conducted measurements within seven days prior to delivery on a sample of 174 fetuses. They conclude that precision remains an unresolved issue in the case of macrosomia. Most methods are based on regression analysis. Hadlock method was published in 1985 and it was based on the aforementioned mathematical calculation.

A study published in 2020 compared accuracy of 22 fetal weight estimation methods among pregnant women with GDM. Not surprisingly, a conclusion was that none of the compared methods was sufficiently accurate to determine fetal macrosomia. Moreover, fetal weight ultrasound examination conducted before delivery in order to assess fetal weight proved to be useless for assessing fetal macrosomia. The fetus is still in the growth phase and in three or four weeks, it may gain sufficient weight to become macrosomic. From the clinical standpoint, fetal weight estimates, if accurate, have good predictive value, provided they were conducted within seven days prior to delivery. As already noted, most methods are based on regression models where precision at the ends of the distribution curve is reduced. Most methods take into account the size of individual fetal biometric parameters and compare it against the actual weight or newborn weight. It is therefore important that fetal weight estimation is conducted as closest as possible to the time of delivery to avoid uncertainty as to what extent fetal weight has increased since the last estimation. Considering the margin of error, deviations of ten or more percent from the actual newborn weight can have an impact on making an incorrect clinical decision, at least in retrospect.

It is not always easy to perform fetal biometrics by ultrasound. Sometimes the head positioned deep into the pelvis does not offer the required cross-section, and sometimes breech presentation makes estimation difficult as well. In larger pregnancies, especially in the case of ruptured amniotic sac, ultrasound estimation of fetal weight is extremely difficult. Furthermore, ultrasound estimation of fetal weight also depends on the expertise of the clinician.

Fetal biometric parameters for ultrasound fetal weight estimation

Predictive values of individual biometric parameters in estimating pregnancy duration are different from those used in estimating fetal weight. Biparietal diameter, head and abdomen circumferences, and thigh bone length are parameters that have been confirmed through numerous studies and are still used as a standard for both gestational age and fetal weight estimates. However, the predictive value also depends on the duration of gestation.

Biparietal diameter

It has been observed that, after the first trimester, the combination of head circumference and femur length had a similar diagnostic value as the combination with other parameters in estimating gestational age. Adding other parameters did not contribute to the accuracy of the calculation.

Biparietal diameter is used in almost all guidelines and most gestation and fetal weight estimation methods. However, it is not as reliable in the advanced stage of pregnancy, mostly due to different head shapes.
Measurement of BPD for purposes of fetal weight estimation can only make sense if a standardized ultrasound section of the fetal head in which it is measured has been determined. Moreover, it is important to conduct correct measurements because of comparison with measurements during pregnancy and other comparisons, such as the need for a second opinion, scientific research, and similar. The plane intersecting the thalamus and the third brain ventricle offers the most credible section. Calvary bones must be placed in a way to present a highly symmetric image, i.e., the left and right sides should reflect as in a mirror. Hadlock measured BPD between the outer sides of the skull bones\textsuperscript{9}.

Some clinicians use the BPD to OFD (occipitofrontal diameter) ratio. When such a ratio is multiplied by 100, the so-called cephalic index is obtained\textsuperscript{9}. It is not much in use in everyday practice but is often used in oligohydramnios, premature rupture of fetal membranes, breech presentation, or rupture of neural tubes that can change the appearance of the fetal head, e.g., in the case of dolichocephaly. Hadlock et al. showed that in most normally shaped fetal heads, the range of cephalic index was approximately one standard deviation. When this index is close to or outside the limits of one standard deviation, the authors concluded that BPD was not a reliable biometric parameter and that preference should be given to head circumference\textsuperscript{9}.

**Head circumference**

Fetal head circumference is used routinely in estimating the duration of pregnancy and fetal weight. Similar to BPD, it is used after the 14\textsuperscript{th} week of pregnancy. Although BPD is the most widely used biometric parameter in fetal weight estimation and is an integral part of most methods for its calculation, certain studies have shown that head circumference is a more reliable parameter than BPD\textsuperscript{10}. These studies were published at the very beginnings of combining different biometric parameters of the fetus for pregnancy duration and fetal weight estimation, encouraged by a new method of looking inside the uterus, i.e., ultrasound. Some recent research denies such an advantage, leaving BPD as an irreplaceable parameter\textsuperscript{11}.

The plane in which to measure head circumference according to standard instructions is similar to that in which we measure BPD. It is a presentation of the thalamus and third ventricle, but also the ‘cavum septum pelucidum’ from the front, and separation of the tentorium from the back. Head circumference is largest in such a cross-section. The appearance of the head should also, similar to when measuring BPD, resemble a mirror image of one half of the head. In this way, we can be assured that the section is not oblique, and that the measurement of fetal head circumference is more reliable. Devices in use nowadays enable circumference measuring by placing an ellipse around the fetal head. The ellipse must be positioned along with the head bones. Position along the scalp will falsely increase the circumference.

**Femur length**

Although seemingly simple, the process of femur length determination must observe certain standardized rules to obtain a reliable measure. Femur length is a rather common biometric parameter used by clinicians on estimating pregnancy duration, especially in uncontrolled or insufficiently controlled pregnancies. Variations in the femur length, aside from racial differences perceivable from the beginning of the second trimester\textsuperscript{12}, are based on biological differences, and it has been observed that femur length correlates with certain genetic disorders and malformations\textsuperscript{13}. The femur to be measured must be as close to the ultrasound probe as possible, in the most possible parallel plane. The femoral head or large trochanter on the proximal part, and the condyle on the distal part must be shown. This achieves greater reliability of the measure, i.e., false overestimation and underestimation of length. The ossified part of the bone should be measured but without the femur head. The cursor should be positioned at the cartilage and bone junction.

**Abdominal circumference**

Abdominal circumference is a biometric parameter that, given its wide range, has greater reliability in estimating fetal weight than gestational age, especially in the case of diabetes and pregnancy\textsuperscript{14,15}. It represents an important biometric parameter in determining gestational age in fetuses with head or leg malformation, given dubious reliability of other common parameters. Apart from pregnant woman and fetus diseases, the range of abdominal circumference measures is influenced by natural variations, as well
as by the expertise of the physician performing the ultrasound examination and by the techniques of display and measurement. Due to the irregular, asymmetric appearance, and poorer echogenicity of the structures in the standardized cross-section, it is not easy to measure abdomen circumference.

The plane to perform the measurement implies presentation of the fetal liver in the largest diameter and presentation of the left and right branches of the portal vein, along with the shortest umbilical part of the left portal vein. Fetal ribs should be shown symmetrically, with the maximum possible vertical cross-section of the fetal abdomen. Cursors or ellipses are placed to include the skin, and not the bony part (different from head circumference measurement).

**Other fetal biometric parameters**

Aside from the above indicated, other biometric parameters were used in the attempts to find a reliable fetal weight estimation method. Given that in the past approximately 30 methods for calculating fetal weight have been developed, it was not surprising to find that none is sufficiently accurate in every obstetric situation. Ultrasound devices in their software generally have several methods and their modifications. Most common are Shepard’s modification of Warsof’s method and Hadlock’s method2.

Considering that Hadlock’s method, along with its modifications, is one of the most widely used in fetal weight estimation, it has been the object of numerous studies. Many researchers, endeavoring to contribute to the accuracy of the method, use individual biometric parameters such as thigh circumference, and some have attempted to calculate fetal weight using artificial intelligence.

**Comparison of methods for ultrasound fetal weight estimation**

It has been shown that fetal weight estimation also depends on the ultrasound measurement technique. In certain studies, using a three-dimensional (3D) ultrasound, a lower incidence of clinically relevant errors in fetal weight estimation was found, whereas others showed opposite results. Given the conflicting data, and the fact that performing 3D ultrasound examinations in the everyday practice of fetal weight estimation is rather complicated (requires more time, knowledge, experience, and more expensive equipment), 3D ultrasound examinations are currently justified only in complex cases or for research purposes.

In consideration of the entirety of the research on fetal weight estimation published thus far, a conclusion would be that a perfectly precise method does not exist. These are more likely to be methods that have shown sufficient precision in the given samples, either in pregnancies in the general population or pregnancies with specific maternal and/or fetal conditions.

Some of the known factors that can interfere with the method accuracy are variations due to gestational age. Most methods are more accurate in estimating fetal weight of term fetuses, which is rather logical when deviation from the actual weight (newborn as a gold standard) includes time from the moment of weight estimation to actual delivery. By designing population-based centile growth curves, such increase in fetal weight is known and available, and empirical calculations based on tables with fetal weight percentiles concerning the duration of pregnancy are often applied, mainly for weekly level of weight gain. Such tables are mainly built on the general population in order to be exhaustive, but the sample described is not necessarily optimal to apply in other conditions. Examples are pregnancies with diabetes. Aside from the greater availability of glucose, which affects fetal growth, changes also occur in insulin signaling and growth factor receptor stimulation.

Most studies related to modification of fetal weight estimation methods used the period from the last ultrasound estimation to delivery longer than seven days as an exclusive criterion. Such logic, as described above, ensured minimum deviation from the actual weight. This is necessary in the coefficient calculation of modified methods in order to minimize deviations from the actual, i.e., newborn weight.

On assessing accuracy of the fetal weight calculation method, all relevant factors that may influence distortion of the test results must be taken into account.

One of the factors contributing most to a possible error is gestational age. It has been perceived that wide ranges in gestational age, in a sample that endeavors to cover all weeks of pregnancy, contribute to reduction in method accuracy. In a study published in 2004, Kurmanavicius et al. compared accuracy of several methods (Campbell, Wilkin, Shepard, Merz, and two Hadlock’s) used in 5612 fetuses. It was concluded that both Hadlock’s methods were more precise.
than the others, although not sufficiently enough to be absolutely accepted, especially for fetuses of small or large weight. Furthermore, there was a difference in the measurement results of individual physicians, and it was concluded that such a difference might cause reduced accuracy. This conclusion leads us to another factor influencing the accuracy of fetal weight calculation, i.e., experience, knowledge and skill of the physician conducting the examination.

Another factor to consider when investigating the accuracy of fetal weight calculation methods are the two ends of fetal weight. Wide ranges of fetal weight reduce the method precision. A possible answer lies in the amount of a particular type of tissue (e.g., adipose tissue), and not just the entire size of a particular parameter. The average deviation error from newborn weight is greater in the case of fetal growth below 10th and above 90th centile.

Samples involving fetuses with greater body weight show more frequent errors in underestimating fetal weight than in the actual newborn weight as the gold standard and vice versa, in the case of lower weight fetuses, errors of weight overestimation are more frequent. There are conflicting studies. Coomarasamy et al. showed that in the case of a positive finding, macrosomia was more present than in the case of a negative finding. This meta-analysis included 63 other studies. A total of 19117 pregnant women were included in the analysis. The authors conclude that in the case of a positive test (predicting macrosomia), it is more reliable to assume that macrosomia is present than to rule it out in the case of a negative test. In other words, considering the possibility of error and data on the continuation of fetal growth from the last estimation to delivery, in borderline findings, macrosomia should be considered as present.

There are other factors that make measured biometric parameter less reliable, mainly the small or greater amniotic fluid volume. The fetus position also contributes to the calculation accuracy. According to certain studies, thickness of the maternal subcutaneous adipose tissue also influences the process of measuring or displaying biometric parameters.

Comparison with the results of research conducted on a multiethnic population should be taken with caution. There are certain differences in the size of individual biometric parameters and caution is envisaged while interpreting the results within a particular study. However, comparing errors with other studies does not necessarily lead to the same alteration of results.

From the relevant literature, it transpires that the majority of researchers use the following methods while comparing different methods of fetal weight estimation: average deviation of estimated weight from actual newborn weight, error frequency rate according to percentage classes (up to 5%, up to 10%, and over 10%), comparing areas under curve (ROC analysis), sensitivity and specificity of the test (along with predictive values).

Combs et al. compared 31 fetal macrosomia estimation methods in the population of diabetic pregnant women and pregnancies longer than 36 weeks, along with measurements performed within two weeks from delivery. The result showed that the area under curve indicated a range from 0.836 to 0.897 and there was no difference in the size of the area under curve among any of the methods. Hadlock's method, depending on its modification, in this study presented an area under curve from 0.87 to 0.89. These areas are consistent with the results of other researchers and are considered to have excellent discrimination. This study included 119 pregnant women with gestational diabetes and 46 with preexisting diabetes. The macrosomia incidence was 30%. Combs et al. concluded that no method prevailed. An example is the Vintzileos method. In the aforementioned study, the systemic error of the method was smallest, but the discriminatory value (determined by the size of the area under curve) was at the bottom of the scale with respect to the absolute error. The opposite result was also found. Certain methods (Thurnau, Weinberger, Woo, and Scot) had such a large systemic and absolute error that the authors concluded that their clinical usability should be reconsidered. The best-ranked method (Ott) had a positive predictive value of 81%. A moderate increase of sensitivity to 90% resulted in 42% of false-positive results, which was rather poor. Moreover, the latter method had a frequency of 32% of errors over 10% deviation from the newborn weight (absolute error). In 7% of these, deviations were greater than 30% of the newborn weight.

Shmueli et al. compared different estimation methods for fetal weight or macrosomia. In their study, the Hadlock 2-method for predicting macrosomia yielded a sensitivity of 77.55% and specificity of 87.12%, which was incidentally the best ratio...
compared to other methods. The positive predictive value was 38%, whereas the negative predictive value was 97.44%. The accuracy was 86.24%. The study had several limitations common to retrospective studies. The sample was a diabetic population, regardless of diabetes type (a total of 1060 subjects). Fetal weight estimates were conducted within seven days to delivery.

We can affirm that tests have a greater ability to rule out macrosomia rather than rule it in. This information is important, considering the research by Bryant et al., which showed that it was necessary to perform 155 to 588 cesarean sections to prevent a single case of injury to brachial plexus of a newborn. In this study, comparison of 10 models for fetal weight estimation resulted in the mean sensitivity of all methods of 51.55% (range from 20.62% to 81.44% in Shepard's method). The specificity was, as expected, higher and amounted to 92.53% (range from 76.64% to 97.09%). The positive and negative predictive value for all methods was 41.44% and 95.01%, respectively. The accuracy of all methods combined was 88.87%, ranging from 77.08% to 91.13%. Analysis of absolute error frequency over 10% in the deviation from the newborn weight showed variable results. Hadlock models ranged from approximately 43% to 82% of estimates within a 10% deviation, whereas other methods presented a frequency of approximately 72% to 80%. In this study, based on a mathematical model for calculating Euclidean distances, Hadlock's method that includes BPD, AC and FL was declared as the second best one.

Chauhan et al. conducted a retrospective research in the general population where ultrasound examinations were performed in term pregnancies, i.e., fetal weight estimation was conducted within three days before delivery. The sensitivity of the methods applied ranged from 25% to 42%, and the specificity from 93% to 98%. The percentage of correctly diagnosed cases of macrosomia was 76.5%. A few years later, the same authors published a study comparing the precision of the method, regardless of macrosomia. In higher fetal weights, the precision was lower, whereas the overall sensitivity was 71%, specificity 92%, and accuracy of diagnosing macrosomia 55%. The Chauhan et al. research from 2005 analyzed the ability of different methods from 14 studies related to the ultrasound method of detecting macrosomia. The results yielded a sensitivity ranging from 12% to 75% and specificity from 68% to 99%, with greater accuracy (17% to 79%) found in a population with a higher frequency of macrosomia. Such a result is not surprising because the frequency of higher macrosomia contributes to the probability of correct prediction, i.e., that there is a bigger share of those who will later be proven as correct. The aforementioned study indicates that Shepard's method yields only 16% to 32% of the macrosomia estimation accuracy.

There are other studies that report a wide range of sensitivity of the fetal weight estimation models used. In one study conducted by American authors, such range spanned from 25% to 75%. The authors used their model and took into account other parameters such as gestational age, parity, maternal height, weight, and maternal weight gain in the last trimester of the pregnancy. If the sensitivity of 80% was taken, the weight for predicting macrosomia was 3550 grams. The authors were satisfied with this result and considered it as a sufficiently precise model. Furthermore, the research was conducted in the general population, white race, and pregnancies without complications. The practical everyday use of this model was questioned, and most gynecologists did not consider the weight of 3550 grams as a limit for terminating pregnancy by cesarean section to avoid the risk of complications (naturally, when there are no other obstetric indications for cesarean section such as significant pelvic narrowing).

Israeli authors conducted a study aimed at detection of macrosomia in a population of diabetic pregnant women, control group, and pregnant women with fetuses suspected as large for gestational age (LGA). They showed, in the overall sample, that the sensitivity of the ultrasound method for detecting macrosomia was approximately 55%, specificity approximately 88%, positive predictive value 48%, and negative predictive value 91%. The research by Benson et al. in diabetic pregnant women showed interesting results. They compared several conventional ultrasound methods for accurate fetal weight measuring and determining macrosomia (Warsof, Shepard, and Hadlock). They found that the mean deviation from newborn weight in Hadlock's 2-method was smallest, 2.6±12.2%. For Warsof, the error was 5.9±12.6%. These results indicated a reliable prediction of fetal weight in the population of diabetic pregnant women. Given a 95% confidence interval of 15% to 22% for the error, this would mean that
Another research conducted in the United States included diabetic pregnant women on insulin therapy, where measurements were performed within seven days prior to delivery. Six regression models were used on fetal weight estimation. There were 19.4% of macrosomic children. The results showed that there were 66% of estimates within the 10% error of newborn weight, using abdominal circumference and femur length (ranging from 53.4% to 66.2%). The study showed a 63% frequency for Hadlock's 2-method. According to the authors, if the estimated fetal weight is 4500 grams, there is approximately a 50% chance that the baby is macrosomic. This speaks in favor of the poor precision of the methods applied at higher fetal weights. The chance of 50% for precise diagnosis is at the level of probability by flipping the coin. The study reports sensitivity of 51%, specificity of 96%, positive predictive value of 75%, negative predictive value of 89%, and test accuracy of 88% for Hadlock's 2-method. As the sensitivity increases, the specificity of the test decreases. Thus, a sensitivity of 69% and specificity of only 77% were determined for Campbell's two-parameter method (abdominal circumference and biparietal diameter). The specificity of the test is the ability to distinguish subjects that do not have a property. In the latter case, 77% of fetal weight estimates may be considered as not macrosomic.

In 2010, Hoopmann et al. published a study on comparison of 36 different methods and average errors or error frequency, considering a certain percentage category (5%, 10%, 20%, and 30%). This retrospective research was conducted on 350 cases of macrosomic children (4000 grams) from singleton pregnancies and without structural anomalies. The estimations were conducted within seven days prior to delivery, of which 52.3% were performed one day before delivery. The results showed a wide range of errors, from the percentage of weight estimation error to the frequency of estimation of different error percentage classes. Thus, the average error was -7.3% (±8.5%) for Hadlock's 2-method (BPD, AC, FL), -6.5% (±5.8%) for Hansmann's method, and -10.0% (±9.0%) for Hadlock's method (AC, FL). The range was from 0.3% to 62.2%, regardless of the sign. The percentage of errors within 5% for the above methods was 2.9%, 4.6%, and 1.7%, respectively. In the class up to 10%, the frequency was 58.0%, 69.4%, 48.9%, respectively. In the class over 10%, the frequency range was quite wide, spanning from 0 to 95.6% (Hart). The precision range of correctly recognized macrosomia models used in Hoopmann et al. research varies. Of 36 models, only 6 exceeded 60% and 4 exceeded 70%. Of these 4 models, the error frequencies in the class up to 10% were 95.6%, 52%, 50.9%, and 38.3%. Except for the 95.6% (Hart) percentage, other models of comparable precision in macrosomia detection had a significantly lower percentage of weight estimates within deviations in the class up to 10%. Comparison of methods declared as most accurate (Ott method, absolute error of 8.2%, relative -2.9%) in the research by Combs et al. was particularly interesting. Hoopmann et al. showed in their study that this method (Ott) presented an exceptionally large absolute error (12.1%) and a relative error (11.8%). Moreover, it had low accuracy in correct estimation of macrosomia (23%).

Discussion on the accuracy of the fetal weight estimation methods against the biometric parameters applied can get further complicated, as shown in the example of fetal AC. According to certain studies, the best biometric parameter for detecting macrosomia is the circumference of fetal abdomen. More specifically, if the range is within the scale from 35 cm to 38 cm, this parameter can function as an independent predictor of macrosomia. Rosati et al. confirm such a finding, adding that, however, a combination with other parameters is more precise. This is confirmed by the results of other studies of method comparison, in which the combination of BPD, AC, and FL proved to be most accurate. However, measuring the circumference of the abdomen is not simple and it cannot be always easily conducted in a standardized cross-section. Such a cross-section includes the largest liver diameter because changes in fetal growth are closely related to liver size. According to the results of the research on the role of AC in fetal weight estimation or macrosomia, the AC proved to be a good predictor provided it is higher by two to three weeks compared to gestational age. In cases above the 90th percentile, it is suggested to repeat fetal weight estimation close to delivery. In cases below the 90th percentile, the results suggest that repeating ultrasound examination for fetal weight estimation will not contribute to the accuracy of macrosomia estimation. However, a study was conducted in Berlin on the population of pregnant women with gestational diabetes (1914 subjects...
and 4478 ultrasound examinations). Limitations of this retrospective study were the frequency of 10% of women having undergone oral glucose test due to suspicion of high fetal weight between 28 and 32 weeks of gestation and uncertain accuracy of the measured biometric parameters due to the lack of physicians experienced in performing ultrasound examination. The authors concluded that macrosomia prediction would be more accurate if certain risk factors related to pregnant women were added to the ultrasound fetal weight estimation. Other authors support such a conclusion, stating that one should not rely exclusively on ultrasound fetal weight estimation and that other risk factors (maternal anthropometric data and presence of diabetes) or data adjustment can significantly contribute to greater precision in macrosomia estimation.47,48

By comparing 26 methods, Melamed et al. concluded that highly precise methods applied to their sample were those that included three or more biometric parameters. A total of 3705 ultrasound estimations were conducted within three days prior to delivery. Such results were not confirmed by the research conducted by Hoopmann et al., who found that four of the five most accurate methods were based on a model with two biometric parameters (Schillinger, Ferrero, Hansmann, and Merz).43

Although certain authors, based on the results of their own research, consider AC as the most important biometric parameter, this information can ‘cut both ways’. The effect on the error or deviation of the estimated fetal weight in relation to the newborn weight may rise to 3.5% in the case of AC. This information may account for the largest share of systematic or accidental error.50,51

Siemer et al. described 11 fetal weight estimation methods in their study. The research included 1941 pregnancies and there were 211 macrosomic children (4000 grams). The relative and absolute error ranged between 5.2% and 7.5% for the Merz method and 14.5% and 14.9% for the Campbell method, respectively. Two of Hadlock’s methods proved to be most accurate, while Schiled’s method was most accurate when gender difference was taken into account (accuracy of 40% in class up to 5% of deviation from the newborn weight). This study showed, as the Kurmanavicius et al. study did, that most methods tended to underestimate fetal weight, most likely because they did not have macrosomic children as a population by which the coefficients in the regression model could be adjusted.

The prediction of a newborn weight ≥4500 grams is even less accurate. The studies report an absolute error rate of approximately 12.6%, which is more than 8.4% in the case of 4000 grams, unrelated to other risk factors such as diabetes. In another study, in children larger than 4500 grams, the error frequency in the class of over 10% of deviation from newborn weight was more than 50%.45

Ultrasound fetal weight estimation has proven to be the most common fetal weight estimation method. To combine other risk factors may be slightly more complicated to apply in everyday work, but most physicians account for such risk factors empirically. Therefore, respective studies (although some already exist) are particularly important for quantification of such models. This would help in making decision based on individual results obtained through recognized scientific methods of their analysis.

Another empirical conclusion refers to the need for serial measurements, especially of those fetuses whose preterm AC is above the 90th percentile for gestational age. Nevertheless, some authors dispute the ratio of input and output of such an approach.56

As a contribution to the overabundance of conflicting results, Irish authors report the results of fetal AC analysis in predicting normal fetal weight (between the 10th and 90th percentile). The study explored the possibility of avoiding unnecessary repeating ultrasound examinations. The results showed that fetuses with AC at the 50th, 75th and 90th percentile had 100%, 97%, and 96% positive predictive value, respectively.

Furthermore, speaking of ultrasound techniques, some studies applied a 3D technique to test the accuracy of predicting large fetal weights. The reason behind such an approach was the fact that 2D ultrasound did not take volume into account. Ultrasound fetal weight estimation does not involve tissue composition and therefore children with almost identical biometric parameters such as BPD, AC, and FL may present differences in newborn weight.

Concerning the latter, some attempts endeavored to increase the fetal weight estimation accuracy by taking into account subcutaneous adipose tissue. Such data would be particularly useful in a fetal population exposed to higher glucose concentrations during pregnancy. The results suggest that accelerated fetal
growth can be diagnosed based on subcutaneous adipose tissue measurements. According to the results, it is estimated that adipose tissue of the newborn has a share of approximately 14% in the total weight of the newborn, but also that such deviations may vary up to 46%.

A smaller study by Italian authors encompassed the results of three studies conducted on a total of 287 fetuses. The study included ultrasound measurements of the soft tissue of the abdomen and thigh of the fetus. This analysis, for all models combined, yielded the 80% success rate in detecting macrosomia. In 2012, a group of authors concluded that 2D ultrasound had no further use in fetal weight prediction, i.e., that precision of the methods using 2D imaging technique had reached its peak and that different approaches should be applied in the ultrasound fetal weight estimation methods. Incidentally, two years later, they published an article using a 2D ultrasound technique to assess fetal weight.

A group of Austrian authors from Salzburg compared fetal weights and analyzed estimation models using a 2D technique (Hansmann, Merz, and Hadlock methods) and 3D technique (Schild, Liang, and Chang methods). Schild's method using the 3D technique showed a moderately higher precision, while the other methods in the 3D view did not prove to be better in comparison to the conventional methods using the 2D technique. It has already been said that 2D measurement of fetal adipose tissue, although well correlated with newborn skin folds, does not significantly contribute to better macrosomia prediction. A clinically important question arises as to whether ultrasound estimation of fetal weight should be used as a screening method to detect shoulder dystocia (and the consequences it carries). There are conflicting data on fetal size and shoulder dystocia. Although there are conflicting data on fetal size, most guidelines are not based solely on the results of fetal weight estimation but determine maternal diabetes as a risk factor, along with fetal weight ≥4500 grams. The study by Hoopmann et al., as described above, presented a combined detection ability of all methods of only 22% for newborns ≥4500 grams. In that study, Birnholz's method had the highest detection rate of 83%, with 46% of false-positive results. The second best was the Hadlock method (Hadlock IV), with 74.5% success in detection and 31% of false-positive results.

**Conclusion**

Considering the aforementioned deviations in the precision of individual fetal weight estimation methods, presented as generally different models, it can be stated that the accurate fetal weight estimation and prediction of macrosomia is an ongoing discussion. Moreover, all indications point to the fact that one should not rely solely on ultrasound fetal weight estimation but should look for new screening models. Such models may include various biometric parameters, ultrasound techniques, risk factors such as individual diseases, anthropometric data, and similar. In the case of a modification of the methods for 2D ultrasound detection, a population that serves as a gold standard in modifying the coefficients of different regression models should be considered.

The challenge of drafting a meaningful conclusion arises when taking into consideration considerably wide ranges of fetal weight estimation accuracy, as discussed in the text above. And we are not speaking only of the methods applied to ultrasound estimation. If we exclude the possibility of combining ultrasound techniques, different types and numbers of biometric parameters, anthropometric data of pregnant women, and other risk factors, the question of comparing data describing the fetal weight estimating methods remains barren. Such comparisons, unfortunately, are not standardized. Some authors used the mean error value, mean error percentage, mean deviation of the estimated weight from the newborn weight, frequency of those measurements considered as clinically accurate (5% and 10% deviation classes), areas under curves, test sensitivity and specificity, predictive values and finally, test accuracy. Other authors advocate the area under curve (ROC analysis), but it is questionable how much such a result can contribute to clinical decision. A sensitivity of 80% has already been mentioned (indicating an 80% probability of having a macrosomic child) if fetal weight amounts to 3550 grams. However, such an approach implies an error of at least 450 grams or 12.6%.

The purpose of fetal weight estimation is to obtain useful data aiming at terminating the pregnancy vaginally (except for certain indications for surgical delivery), with a low risk of complications, based on the elevated weight of the child. On the other hand, there is a need to identify the weight of a fetus posing an unacceptably high risk of injury at vaginal delivery.
Fetal weight estimation methods are overly general. Methods in their nature imply an increase in weight with respect to an increase in the 2D value of the measured parameter. Weight is a function of volume, not just an exponential function. Furthermore, difference in the density of individual tissue types and their representation (not necessarily volume) must not be neglected as they also affect the weight.

Protocols for identifying fetal macrosomia will certainly continue, as long as there are needs, opportunities, and sufficiently motivated researchers who want to contribute to the health of unborn but full-fledged human beings.

References

1. Sabolović Rudman S, Djaković I, Gall V, Djaković Ž, Košec V. Pregnancy outcome in gestational diabetes compared to body mass index. Acta Clin Croat. 2019 Mar;58(1):37-41. doi: 10.20471/acc.2019.58.01.05.

2. Hadlock FP, Harrist RB, Sharman RS. Estimation of fetal weight with the use of head, body and femur measurements: a prospective study. Am J Obstet Gynecol. 1985;151:333-7. doi: 10.1016/0002-9378(85)90298-4.

3. Esinler D, Bircan O, Esin S, Sahin EG, Kandemir O, Yalvac S. Finding the best formula to predict the fetal weight: comparison of 18 formulas. Gynecol Obstet Invest. 2015;80(2):78-84. doi: 10.1159/000365814.

4. 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 Med. 2012 Dec;33(7):E62-E67. doi: 10.1055/s-0031-1281833.

5. Cesnaite G, Domza G, Ramasauskaite D, Volochovic J. The accuracy of 22 fetal weight estimation formulas in diabetic pregnancies. Fetal Diagn Ther. 2020;47(1):47-50. doi: 10.1159/000500452.

6. Hadlock FP, Deter RL, Harrist RB, Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. Radiology. 1984 Aug;152(2):497-501. doi: 10.1148/radiology.152.2.6739822.

7. Persson PH, Waeber BM. Reliability of ultrasound fetometry in estimating gestational age in the second trimester. Acta Obstet Gynecol Scand. 1986;65(5):481-3. doi: 10.3109/00016348609157390.

8. Papageorghiou AT, Kemp B, Stones W, Ohuma EO, Kennedy SH, Purwar M, et al. International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Ultrasound-based gestational-age estimation in late pregnancy. Ultrasound Obstet Gynecol. 2016 Dec;48(6):719-26. doi: 10.1002/uog.15894.

9. Hadlock FP, Deter RL, Carpenter RJ, Park SK. Estimating fetal age: effect of head shape on BPD. AJR Am J Roentgenol. 1981 Jul;137(1):83-5. doi: 10.2214/ajr.137.1.83.
diabetic pregnancies. J Anat. 2009 Jul;215(1):60-8. doi: 10.1111/j.1469-7580.2008.01035.x.

24. Kurmanavicius J, Burkhardt T, Wisser J, Huch R. Ultrasonographic fetal weight estimation: accuracy of formulas and accuracy of examiners by birth weight from 500 to 5000 g. J Perinat Med. 2004;32(2):155-61. doi: 10.1515/JPM.2004.028.

25. Dudley NJ. A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol. 2005 Jan;25(1):80-9. doi: 10.1002/uog.1751.

26. Committee on Practice Bulletins – Obstetrics. Macrosomia: ACOG Practice Bulletin, Number 216. Obstet Gynecol. 2020 Jan;135(1):e18-e35. doi: 10.1097/AOG.0000000000003606.

27. Coomarasamy A, Connock M, Thornton J, Khan KS. Accuracy of ultrasound biometry in the prediction of macrosomia: a systematic quantitative review. BJOG. 2005 Nov;112(11):1461-6. doi: 10.1111/j.1471-0528.2005.00702.x.

28. Drassinower D, Timofeev J, Huang CC, Benson JE, Driggers RW, Landy HJ. Accuracy of clinically estimated fetal weight in pregnancies complicated by diabetes mellitus and obesity. Am J Perinatol. 2014 Jan;31(1):31-7. doi: 10.1055/s-0033-1334450.

29. Tsai PJ, Loichinger M, Zahal I. Obesity and the challenges of ultrasound abnormality diagnosis. Best Pract Res Clin Obstet Gynaecol. 2015 Apr;29(3):320-7. doi: 10.1016/j.bpobgyn.2014.08.011.

30. Maxwell C, Glanc P Imaging and obesity: a perspective during pregnancy. AJR Am J Roentgenol. 2011 Feb;196(2):311-9. doi: 10.2214/AJR.10.5849.

31. Buck Louis GM, Grewal J, Albert PS, Sciscione A, Wing DA, Grobman WA, et al. Racial/ethnic standards for fetal growth: the NICHD Fetal Growth Studies. Am J Obstet Gynecol. 2015 Oct;213(4):449.e1-449.e41. doi: 10.1016/j.ajog.2015.08.032.

32. Combs CA, Rosenn B, Miodovnik M, Siddiqui TA. Sonographic EFW and macrosomia: is there an optimum formula to predict diabetic fetal macrosomia? J Matern Fetal Med. 2000 Jan-Feb;9(1):55-61. doi: 10.1080/1524000000897945.

33. Hosmer DW, Lemeshow S, Sturdivant RX. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, Sturdivant RX, editors. Applied Logistic Regression. 3rd edn. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2013. p. 177.

34. Shmueli A, Salman L, Hadar E, Aviram A, Bardin R, Ashwal E, et al. Sonographic prediction of macrosomia in pregnancies complicated by maternal diabetes: finding the best formula. Arch Gynecol Obstet. 2019 Jan;299(1):97-103. doi: 10.1007/s00404-018-4934-y.

35. Bryant DR, Leonard MR, Landwehr JB, Bottoms SF. Limited usefulness of fetal weight in predicting neonatal brachial plexus injury. Am J Obstet Gynecol. 1998;179:686-9. doi: 10.1016/s0002-9378(98)70065-1.

36. Chauhan SP, Perry KG Jr, Magann EF, Lutton TC, Meydrech EF, Morrison JC. Intrapartum assessment of birth weight at term: clinical versus sonographic models using from one to four fetal parameters. J Matern Fetal Invest. 1994;4:263-7.
Diagnosis of macrosomia in gestational diabetes

B. Lovrić et al.

variability. J Clin Ultrasound. 1993 Oct;21(8):515-9. doi: 10.1002/jcu.1870210808.

Sarmandal P, Bailey SM, Grant JM. A comparison of three methods of assessing inter-observer variation applied to ultrasonic fetal measurement in the third trimester. Br J Obstet Gynaecol. 1989 Nov;96(11):1261-5. doi: 10.1111/j.1471-0528.1989.tb03222.x.

Smulian JC, Ranzini AC, Ananth CV, Rosenberg JC, Vintzileos AM. Comparison of three sonographic circumference measurement techniques to predict birth weight. Obstet Gynecol. 1999 May;93(5 Pt 1):692-6. doi: 10.1016/s0029-7844(98)00517-1.

Sarmandal P, Bailey SM, Grant JM. A comparison of three methods of assessing inter-observer variation applied to ultrasonic fetal measurement in the third trimester. Br J Obstet Gynaecol. 1989 Nov;96(11):1261-5. doi: 10.1111/j.1471-0528.1989.tb03222.x.

52. Smulian JC, Ranzini AC, Ananth CV, Rosenberg JC, Vintzileos AM. Comparison of three sonographic circumference measurement techniques to predict birth weight. Obstet Gynecol. 1999 May;93(5 Pt 1):692-6. doi: 10.1016/s0029-7844(98)00517-1.

53. Siemer J, Egger N, Hart N, Meurer B, Müller A, Darhe O, et al. Fetal weight estimation by ultrasound: comparison of 11 different formulae and examiners with differing skill levels. Ultraschall Med. 2008 Apr;29(2):159-64. doi: 10.1055/s-2007-963165.

54. Alsulyman OM, Ouzounian JG, Kjos SL. The accuracy of intrapartum ultrasonographic fetal weight estimation in diabetic pregnancies. Am J Obstet Gynecol. 1997 Sep;177(3):503-6. doi: 10.1016/s0002-9378(97)70136-4.

55. Benacerraf BR, Gelman R, Frigoletto FD Jr. Sonographically estimated fetal weights: accuracy and limitation. Am J Obstet Gynecol. 1988 Nov;159(5):1118-21. doi: 10.1016/0002-9378(88)90425-5.

56. Tarca AL, Hernandez-Andrade E, Ahn H, Garcia M, Xu Z, Korzeniewski SJ, et al. Single and serial fetal biometry to detect preterm and term small- and large-for-gestational-age neonates: a longitudinal cohort study. PLoS One. 2016 Nov 1;11(11):e0164161. doi: 10.1371/journal.pone.0164161.

57. Neff KJ, Walsh C, Kinsey B, Daly S. Serial fetal abdominal circumference measurements in predicting normal birth weight in gestational diabetes mellitus. Eur J Obstet Gynecol Reprod Biol. 2013 Sep;170(1):106-10. doi: 10.1016/j.ejogrb.2013.05.023.

58. Farah N, Stuart B, Donnelly V, Rafferty G, Turner M. What is the value of ultrasound soft tissue measurements in the prediction of abnormal fetal growth? J Obstet Gynaecol. 2009 Aug;29(6):457-63. doi: 10.1080/14436109030003209.

59. Bernstein IM, Catalano PM. Influence of fetal fat on the ultrasound estimation of fetal weight in diabetic mothers. Obstet Gynecol. 1992 Apr;79(4):561-3.

60. Maruotti GM, Saccone G, Martineilli P. Third trimester ultrasound soft-tissue measurements accurately predicts macrosomia. J Matern Fetal Neonatal Med. 2017 Apr;30(8):972-6. doi: 10.1080/14767058.2016.1193144.

61. Kehl S, Schmidt U, Spaich S, Schild RL, Sütterlin M, Siemer J. What are the limits of accuracy in fetal weight estimation with conventional biometry in two-dimensional ultrasound? A novel parturient study. Ultrasound Obstet Gynecol. 2012 May;39(5):543-8. doi: 10.1002/uog.10094.

62. Hasenoehrl G, Pohlhammer A, Gruber R, Staudach A, Steiner H. Fetal weight estimation by 2D and 3D ultrasound: comparison of six formulas. Ultraschall Med. 2009 Dec;30(6):585-90. doi: 10.1055/s-0028-1109185.

63. Naef RW 3rd, Martin JN Jr. Emergent management of shoulder dystocia. Obstet Gynecol Clin North Am. 1995 Jun;22(2):247-59.

64. Chauhan SP, Gherman R, Hendrix NW, Bingham JM, Hayes E. Shoulder dystocia: comparison of the ACOG practice bulletin with another national guideline. Am J Perinatol. 2010 Feb;27(2):129-36. doi: 10.1055/s-0029-1224864.

Sažetak

ULTRAZVUČNA DIJAGNOSTIKA MAKROSOMIJE KOD ŽENA S GESTACIJSKIM DIJABETESOM – PREGLED LITERATURE

B. Lovrić, S. Šijanović, I. Zmijanović, G. Jurić i J. Juras

Trudnoće opterećene gestacijskim dijabetesom (GDM) vjerojatnije će, u usporedbi s općom populacijom, završiti rođenjem makrosomskog djeteta, gdje je učestalost operativnog dovršenja trudnoće češća, popraćeno s više komplikacija i oštećenja majke i djeteta. Potreba za izračunavanjem fetalne mase neposredno prije porođaja dovela je do razvoja brojnih metoda u svrhu veće preciznosti procjene. Pregledali smo literaturu od 1980. do 2020. godine rabeći izraze makrosomija, ultrazvučna procjena, gestacijski dijabetes, i upotrijebili smo relevantne članke u pripremi ovog članka. Najčešće primjenjivane metode temelje se na dvodimenzionalnim ultrazvučnim mjerenjima pojedinih fetalnih biometrijskih parametara i njihovoj kombinaciji u matematičkom regresijskom modelu. Neke su metode uključivale dodavanje drugih stanja majke i djeteta kako bi se povećala pouzdanost metode u prepoznavanju makrosomije. U svakodnevnom radu, osobito kod trudnica koje pate od GDM-a, potrebno je imati pouzdane podatke o procijenjenoj fetalnoj težini prije donošenja ispravne kliničke odluke o načinu dovršenja trudnoće. S tim u vezi donosimo pregled literature koja se odnosi na procjenu fetalne makrosomije, naročito u žena s GDM-om.

Ključne riječi: Gestacijski dijabetes mellitus; Fetalna masa; Ultrazvučak; Procjena; Makrosomija