Gevaerd Martins Juliana (Orcid ID: 0000-0003-1327-8979)
Papageorghiou Aris (Orcid ID: 0000-0001-8143-2232)

The influence of maternal body mass index on interobserver variability of fetal ultrasound biometry and amniotic fluid assessment in late pregnancy

J.G. Martins[a], T. Kawakita[a], M. Gurganus[a], D. Baraki[a], P Jain[a], A.T. Papageorghiou[b], A.Z. Abuhamad[a]

[a]Eastern Virginia Medical School, Maternal Fetal Medicine, Department of Obstetrics and Gynecology, USA

[b]St George's University of London, UK and University of Oxford, UK, Maternal Fetal Medicine, Department of Obstetrics and Gynecology, UK

Corresponding author: Juliana Gevaerd Martins

Email: martinjg@evms.edu

Mailing address: 825 Fairfax Avenue, Norfolk, VA 23507

Short title: BMI and fetal biometry interobserver variability

Key words: maternal obesity, fetal biometry, ultrasound interobserver variability, third trimester

What are the novel findings of this work?

This is an unprecedented large prospective longitudinal study that assessed the interobserver variability of fetal ultrasound biometry and amniotic fluid measurements stratified according to maternal BMI categories in late pregnancy.

What are the clinical implications of this work?

Late pregnancy ultrasound has an important role in the detection of fetal growth abnormalities and to guide clinical decisions. Data suggests that ultrasound examination is more technically difficult

in obese patients. Our study shows that obesity does not seem to negatively impact the ultrasound reproducibility of fetal biometric parameters.

Abstract

Objectives: Determine interobserver reproducibility of fetal ultrasound biometry and amniotic fluid assessment in the third trimester according to BMI categories.

Methods: Prospective cohort of women with singleton gestations beyond 34 weeks, recruited into 4 groups according to BMI categories: normal, overweight, obese and morbid obese. Multiple pregnancies, diabetes, growth and fetal abnormalities were excluded. Biometric and fluid measurements were obtained by two experienced physicians/sonographers, blinded for gestational age and each other's measurements. Differences between observers were expressed as gestational age-specific Z-scores. Interobserver correlation coefficient (ICC) and Cronbach's reliability coefficient (CRC) were calculated. Bland-Altman plots were constructed to assess the level of reproducibility.

Results: 110 women were prospectively enrolled (1,320 measurements obtained by 17 sonographers / physicians): 20 had normal BMI (18.2%), 30 were overweight (27.3%), 30 were obese (27.3%) and 30 were morbidly obese (27.3%). Except for MVP (CRC 0.66), all biometric parameters and AFI had high level of reproducibility (CRC 0.84 – 0.93). Among BMI groups: BPD had the highest level of reproducibility (CRC > 0.90); AC and FL reproducibility increased with increasing BMI while MVP decreased. Interobserver differences for biometry fell within the 95% limits of agreement.

Conclusions: Obesity does not seem to negatively impact the reproducibility of fetal biometric parameters when undertaken by experienced physicians/sonographers.

Introduction

Third trimester pregnancy ultrasound has an important role in the management of complicated pregnancies and is frequently used if there is a clinical indication, or routinely in some settings. Accurate estimation of fetal biometry and fetal weight (through a combination of individual parameters) has an important role in the detection of fetal growth abnormalities[1], to monitor fetal growth and to guide clinical decisions[2]. Therefore, accuracy and reproducibility of ultrasound-based fetal biometry in the third trimester is important, as this is the period when growth assessment is most likely to influence clinical decisions regarding timing and delivery mode[2,3]. Systems for quality assurance of ultrasound measurements have been undertaken successfully[4], but these are rarely undertaken at scale in routine clinical practice.

Many previous studies examining reproducibility were limited by small and heterogeneous populations[5], a narrow range of gestational age[5,6], use of old ultrasound equipment[7], limited biometric indices examined[7,8], and varied skills of sonographers performing the ultrasound examinations[5,7-12]. In a recent large prospective longitudinal study performed in low-risk pregnancies under standardized conditions, intra and interobserver variability of fetal biometry measurements throughout the pregnancy were evaluated[2]. The authors concluded that intra and interobserver variabilities, reported as Z-scores, were constant across the gestational age range[2]. However, the study purposely excluded obese pregnant women, meaning we do not have good information in this population[2].

This is important because the obesity epidemic, now estimated to affect over one-third of the United States population, increases pregnancy risks[13,14]. The US National Center for Health Statistic Data from 2015-2016 demonstrated that 41.2% of women of reproductive age are classified as obese and 25.8% of live births in 2015 were to obese mothers[14].

Ultrasound examinations are technically more difficult in obese patients compared to those with normal BMI[11,15]. Few studies have evaluated the ultrasound accuracy for fetal size estimation in the third trimester in obese patients and results have been conflicting, with some[16-18] suggesting that accuracy of sonographic estimation of fetal weight is not influenced by maternal BMI whereas others[19-22] reported decreased accuracy with increasing maternal obesity. Clinical experience certainly suggests that increased BMI is associated with suboptimal ultrasound visualization which makes biometric measurement more difficult, but surprisingly the evidence on the effects of maternal obesity on interobserver variation of such measurement is scant.

The aim of this study is to determine whether there is a relationship between interobserver variability of fetal ultrasound biometry and amniotic fluid in the third trimester and increased maternal BMI.

Methods

This was a prospective observational cohort study performed at the Department of Obstetrics and Gynecology – division of Maternal-Fetal Medicine (MFM), at Eastern Virginia Medical School, United States of America (USA) from May, 2018 to July, 2019. In our hospital third trimester ultrasound is undertaken on the basis of clinical indication. Pregnant women with singleton gestations who presented for an ultrasound examination or for a prenatal care visit at or beyond 34 weeks of gestation were invited to participate in the study. Women were recruited consecutively up to a quota, as our aim was to recruit a roughly equal number of women overweight, obese and morbidly obese groups to facilitate analysis; as well as group of women with normal BMI as a comparator.

Inclusion criteria included maternal age beyond 18 years and BMI above 18.5. Women with multiple pregnancies, presence of pregestational or gestational diabetes, pregnancies complicated by fetal growth abnormalities, structural or genetic abnormalities were excluded to avoid confounding factors related to fetal growth when assessing the reproducibility of the biometric measurements. Pregnancies without reliable dating were also excluded. Each woman was only included once in the study.

Pregnancy dating was established based upon the recommendations of the American College of Obstetricians and Gynecologists (ACOG)[23,24]. The estimated due date (EDD) was assigned according to CRL measurements instead of last menstrual period (LMP) if ultrasound dating before 8 6/7 weeks differed more than 5 days from LMP or if dating from 9 0/7 to 13 6/7 weeks differed from LMP more than 7 days. In the second trimester, the EDD was assigned according to ultrasound fetal biometry instead of LMP if ultrasound dating from 14 0/7 to 15 6/7 weeks differed from LMP more than 7 days or if dating from 16 0/7 to 20 0/7 weeks differed from LMP more than 10 days[23,24]. Pregnancies dated after 20 0/7 weeks were excluded.

The study was approved by the Institution's Review Board (IRB#18-05-FB-0111-EVMS) and written informed consent was obtained from all study subjects.

Eligible women were recruited into 4 groups according to third trimester BMI as follows: normal (BMI of 18.5 – 24.9), overweight (BMI of 25 – 29.9), obese (BMI of 30 -39.9) and morbidly obese (BMI of 40 or greater). We chose to assess the BMI at the time of the examination rather than use the booking BMI, in order to more accurately reflect the maternal habitus at the time of the scan.

All study ultrasound examinations were performed by sonographers certified by the Registered Diagnostic Medical Sonographer (RDMS) and with over 4 years of experience in our ultrasound unit, and physicians with American Institute of Ultrasound in Medicine physician qualifications for the performance of obstetric ultrasound examination (https://www.aium.org/resources/guidelines/obstetric.pdf). The EVMS-MFM ultrasound unit has been continually AIUM accredited for the performance of obstetrical sonography for 11 years. Of note, sonographers and physicians at our center will undertake routine examinations in a large proportion of overweight and obese women in daily practice. Due to the nature of the study, the sonographers and physicians performing the biometric measurements were also considered as study subjects and were consented for the data collection process.

The ultrasound equipment used for the study was the Voluson E6, E8 and E10 (GE Healthcare Ultrasound – Zipf, Austria) with a 4-8 MHz or 6 MHz transabdominal probe. Fetal biometric measurements were acquired based upon recently published imaging criteria for each fetal biometric plane[25]. Minimal pressure was applied on the maternal abdomen during the ultrasound examination to avoid distortion of the measurements.

The corresponding gestational age to the biometric measurement displayed on the screen of the ultrasound machine was blocked from view to minimize bias[26]. Biometric measurements were obtained independently by two sonographers or physicians (referred to as Operators), who were also blinded to gestational age. For each fetus, the first operator performed one complete set of biometric and amniotic fluid measurements. Following an average time lapse of 5 minutes, the second operator repeated all measurements, blinded to the measurements of the first operator. The study imaging protocol included the following: transverse transthalamic plane of the fetal head for the biparietal diameter (BPD) and head circumference (HC), transverse plane of the fetal abdomen for the abdominal circumference (AC), and a longitudinal plane of the femur for the femur length (FL). Amniotic fluid assessment was performed by measuring the amniotic fluid index (AFI) in four quadrants and the maximal vertical pocket (MVP).

The primary outcome of the study was the interobserver reproducibility of fetal ultrasound biometric measurements and amniotic fluid assessment in the third trimester between two experienced imagers according to BMI categories.

Statistical Analysis

Statistical analyses were performed using Stata/IC 16.1 (StataCorp, College Station, TX). A sample size of 20 subjects per BMI group allows the detection of significant differences between groups at a significance level of 0.05 an 80% power.

To test for evidence of non-normality, univariate analysis was performed to examine the distribution of fetal biometric measurements and amniotic fluid assessment taken by each sonographer. Differences between observers for each biometric and amniotic fluid measurement were expressed as gestational age-specific Z-scores. Gestational age-specific Z-scores were calculated by the measurement differences divided by the corresponding standard deviation of that specific fetal measurement for that gestational age[27,28]. Due to lack of normograms that have evaluated means and standard deviations for AFI and MVP in the United States population, Z-scores were not performed for AFI and MVP.

Inter-observer correlation (ICC) coefficients and Cronbach's reliability coefficient were calculated between sonographers for each fetal biometric and amniotic fluid measurement. An ICC value of 1.00 was considered perfect reproducibility between sonographers, while a value > 0.70 was considered very high level of reproducibility. Cronbach's alpha >0.70 was considered to demonstrate good internal consistency. Bland-Altman plots were constructed to assess the level of reproducibility of biometric and amniotic fluid measurements between sonographers using gestational age-specific Z-scores. This allows assessment of systematic bias and 95% limits of agreement. In order to ascertain whether maternal BMI contribute to measurement variability, the corresponding Z-scores were compared using one-way analysis of variance was used[2]. All hypothesis testing was carried out at the 95% significance level, unless otherwise specified, with a p value of <.05 accepted as statistically significant.

Results

Demographics

A total of 110 women were enrolled, resulting in a total of 220 ultrasound studies, 1,320 fetal biometric and amniotic fluid measurements, obtained by 17 different sonographers or physicians. 20 (18.2%) women had a normal BMI, 30 (27.3%) women were overweight, 30 (27.3%) women were obese and 30 (27.3%) women were morbidly obese. The normal BMI group consisted of a lower number of women due to the high prevalence of obesity in our patient population and because the principal reason women with normal BMI were scanned in the third trimester was due to fetal growth restriction, an exclusion criterion. The overweight and obese categories had more women than what was initially determined by the sample size calculation to account for any drop out of subjects, which ended up not happening.

The mean gestational age was 36.7 weeks (± standard deviation 1.5) and mean maternal BMI was 33.5 (± standard deviation 9.0). Maternal age and BMI of women enrolled in the study are displayed in Table 1. Out of the 220 ultrasound studies, 90 were performed by the first author. The remainder of the studies (130) were performed by 16 additional sonographers or physicians. The GE Voluson E6, E8, E10 equipment performed 152, 30 and 38 of the total number of ultrasound studies, respectively.

Fetal Biometric Measurements

Descriptive statistics of the interobserver reproducibility for each fetal biometric (BPD, HC, AC, FL, and EFW) measurement are shown in Table 2. There was very high level of reproducibility for all fetal biometric measurements, with ICC estimates between 0.72 and 0.87. Sonographers displayed consistent reliable measurements for all fetal biometric measurements with reliability coefficient values ranging from 0.84-0.93. The mean Z-score differences were small, ranging from -0.28 to 0.05.

Amniotic Fluid Measurements

Descriptive statistics of the interobserver measurements for amniotic fluid measurements (AFI and MVP) are shown in Table 3. The ICC estimates were 0.65 for AFI and 0.49 for MVP. The sonographers displayed consistent reliable measurements for AFI (reliability coefficient of 0.78) however MVP showed lower levels of reliability (reliability coefficient of 0.66).

Differences by BMI

The reproducibility in fetal biometric measurements and mean Z-score difference according to BMI groups are presented in Table 4. Across all BMI groups, BPD ICC estimates were greater than or equal to 0.80, with reliability coefficients >0.90. Reproducibility of abdominal circumference and femur length measurements appeared to increase with increasing BMI, while the reproducibility of maximum vertical pocket agreement decreased. With exception to the normal BMI group, the maximum vertical pocket measurements showed poor reproducibility between sonographers. However, only Z-scores of femur length were statistically significant (P-values: BPD 0.18; HC 0.67; AC 0.06; FL 0.01; and EFW 0.25).

Figures 1-5 demonstrate Bland - Altman plots for reproducibility between observers for BPD, HC, AC, FL and EFW. In general, the interobserver differences were evenly distributed above and below the zero-difference line, suggesting no systematic errors between observers. Interobserver differences for BPD, HC, AC, FL and EFW primarily fell within the 95% limits of agreement.

Discussion

In our study we evaluated the reproducibility of sonographic fetal biometric parameters in the third trimester, since this is the period when growth assessment is most likely to influence clinical decisions regarding timing and delivery mode[2,3]. Given that obesity is prevalent, associated with many adverse pregnancy outcomes and often leads to suboptimal ultrasound visualization[11,26], our aim was to evaluate the reproducibility of fetal ultrasound biometric measurements in this population.

We found that all fetal biometric measurements are highly reproducible between different observers in normal and increased BMI women in late pregnancy. When comparing Z-score differences, we found that measurement variability of FL was statistically significant, suggesting that measurement variability of FL paradoxically decreased with increasing BMI. Measurement of other biometric parameters, and variability of EFW, were not statistically significant.

The usefulness of a screening test depends on its predictive value, which is affected by its reproducibility[2]. The importance of ascertaining this for sonographic measurements is necessary to improve the accuracy of fetal size and weight estimation[12] and many clinical decisions depend upon accurate and reproducible measurement of fetal biometry.[12] However, despite the alarming increase in obesity rates in the USA and worldwide and the association of obesity with adverse pregnancy outcome, we found no other robust studies that have undertaken assessment of fetal measurement in relation to obesity. The only large prospective longitudinal study that assessed the intra and interobserver variability in fetal biometric measurements was performed in the low risk population under near-optimal conditions, excluding high-risk conditions such as obesity[2].

To our knowledge, this is the first study designed to address this limitation. Strengths of the study are the prospective data collection and a clear, predefined protocol including obese patients according to their BMI category. Unlike most previous studies, which calculated BMI using weight and height measurements from the beginning of the pregnancy[17], we assessed BMI at the time of the scan in order to better reflect the impact of the current BMI on measurement reproducibility and, by extension, how BMI affects these measurements in clinical practice. All standard biometric parameters were evaluated, using standardized measurement techniques with modern and commercially available ultrasound equipment. The study was designed to ensure blinding during measurements and to remove the possibility of sonographers' bias by seeing previous measurements, and/or adjusting to the gestational age displayed on the monitor. In addition, the ultrasound exams were performed by trained sonographers with significant clinical experience, to avoid confounding factors when assessing the reproducibility of the biometric measurements.

Our study demonstrated that all fetal biometric measurements are highly reproducible between different observers in normal and increased BMI women in the third trimester as the overall ICC and reliability coefficient were high for all biometric parameters. Our results are similar to previous studies that evaluated interobserver variability of fetal biometric measurements[2,11,12]. However, our study provides additional information using modern ultrasound equipment, and related to ICC in subclasses of obese pregnancies. Previous studies examining reproducibility of fetal biometry measurements used older ultrasound equipment[7], did not examine all biometric parts[5,17] and had the measurements performed by less experienced operators[11], thus limiting practical applicability. A previous study that compared the impact of maternal weight on intra and interobserver reproducibility of fetal ultrasonography measurements in the third trimester included overweight and obese women in the same group, but only 10% of the women were obese and there was no stratification according to the BMI category[11].

When the ICC and reliability coefficient of fetal biometric measurements were compared between normal and increased BMI groups, only BPD showed excellent reliability across the groups (Table 4). A possible explanation for this finding is that the cranial cavity is an echogenic structure due to its bony composition, typically not compressed by the ultrasound probe. Also, the BPD is obtained by a linear measurement, which results in less variation than measurements taken for curvilinear structures such as the HC and AC[2,12]. As depicted in the Bland and Altman plots (Figure 1 and 2), the BPD mean difference values between two sonographers were close to zero however they showed a slightly increased difference with increasing BMIs. This finding is consistent with previous studies that associate obesity with suboptimal ultrasound visualization[19-22].

The ICC and reliability coefficient for the AC paradoxically increased as BMI increased, with the highest values in the morbidly obese population. This finding is unexpected and may reflect two possible hypotheses: a) more time was spent scanning obese women due to technical difficulties or b) limited sonographic windows are usually observed in obese women which could have improved consistency in measuring biometry. Unfortunately, the study was not designed to test these hypotheses.

AFI and MVP are both techniques commonly used in clinical practice. Studies have previously reported that all sonographic methods of assessing amniotic fluid volume poorly predict adverse pregnancy outcome[29]. Results from a large trial showed that both methods have a high sensitivity to identify normal amniotic fluid volume, but that both over-diagnose amniotic fluid abnormalities, with AFI over-diagnosing oligohydramnios and MVP over-diagnosing polyhydramnios[30]. A previous, large study assessing intraobserver and interobserver reproducibility of measurement of amniotic fluid index (AFI) and MVP (but excluding overweight women) showed that coefficients of variation were similar and limits of agreement were wide for both methods; and

that the choice of method should be dictated by clinical considerations other than method reproducibility[31]. A systematic review of randomized trials[32] determined that MVP should be the method of choice since AFI increases the rate of diagnosis of oligohydramnios, induction of labor and caesarean delivery without improving pregnancy outcomes. Subsequent trials have affirmed these findings[33]. Our study is the first to assess the interobserver variability of AFI and MVP according to BMI categories. Of all measurements in our study, the interobserver differences were highest for AFI and MVP, indicating that fluid assessment is poorly reproducible between observers across all groups. Maximum vertical pocket measurements showed poor reliability and consistency between sonographers.

Our study has limitations. Assuming that the minimum ICC =0.5, an expected ICC = 0.65, alpha = 0.0125, power 0.8, and number of raters = 2, the number of subjects needed would be 230. Assuming a 10% drop-out rate, the number of subjects needed would be 256[34]. Because of the pragmatic design of the study and the fact that multiple sonographers were involved, we planned the study to mimic common conditions in the third trimester. We recognize the sample size limitation however, this represents the first study to systematically look into the effect of obesity in the interobserver variability of fetal biometry in the third trimester. Although the first author (JM) performed ultrasound for almost half of the participants, the remainder of the ultrasound studies were performed by various sonographers. Therefore, the effect of multiple sonographers on the results of statistical analysis is unknown. However, despite the possibility of inter-sonographer differences, we did find high reproducibility of fetal biometric parameters. Our study results may not be applicable to other centers as sonographic training and expertise may impact the internal validity of the study. Although our study shows that AFI and MVP lack interobserver variability in the third trimester, the clinical significance of this finding needs to be further assessed in pregnancies with fluid abnormalities.

In conclusion, ultrasound fetal biometry measurements in women with increased BMI is reproducible when performed by trained sonographers, but amniotic fluid measurements are poorly reproducible. Obesity does not seem to negatively affect the reproducibility of fetal biometric parameters in this controlled environment, when undertaken by sonographers who commonly assess overweigh, obese and morbidly obese women.

Acknowledgement

ATP is supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Melamed N, Yogev Y, Meizner I, Mashiach R, Bardin R, Ben-Haroush A. Sonographic fetal weight estimation. J Ultrasound Med 2009;28:617-629

2. Sarris I, Ioannou C, Chamberlain P, Ohuma E, Roseman F, Hoch L, Altman DG, Papageorghiou AT. Intra- and interobserver variability in fetal ultrasound measurements. Ultrasound Obstet Gynecol 2012; 39: 266-273

3. Lima JC, Miyague AH, Filho FM, Nastri CO, Martins WP. Biometry and fetal weight estimation by two-dimensional and three-dimensional ultrasonography: an intraobserver and interobserver realiability and agreement study. Ultrasound Obstet Gynecol 2012;40:186-193

4. Cavallaro A, Ash ST, Napolitano R, Wanyonyi S, Ohuma E, Molloholli M, Sande J, Sarris I, Ioannou C, Norris T, Donadono V, Carvalho M, Purwar M, Barros F, Jaffer Y, Bertino E, Pang R, Gravett M, Salomon L, Noble J, Altman D, Papageorghiou AT. Quality control of ultrasound for fetal biometry: results from the INTERGROWTH-21st Project. Ultrasound Obstet Gynecol 2018; 52(3):332-339

5. Nakai A, Oya A. Accuracy and reproducibility of ultrasound measurements in obstetric management. Gynecol Obstet Invest 2002; 54:31-36

6. Verburg BO, Mulder P, Hofman A, Jaddoe V, Witteman J, Steegers E. Intra- and interobserver reproducibility study of early fetal growth parameters. Prenat Diagn 2008; 28:323-331

7. Deter RL, Harrist RB, Hadlock FP, Carpenter RJ. Fetal head and abdominal circumferences: evaluation of measurement errors. J Clin Ultrasound 1982;10:357-363

8. Tamura RK, Sabbagha RE. Percentile ranks of sonar fetal abdominal circumferences. Am J Obstet Gynecol 1980;138:475-479

9. Neufeld LM, Wagatsuma Y, Hussain R, Begum M, Frongillo EA. Measurement error for ultrasound fetal biometry performed by paramedics in rural Bangladesh. Ultrasound Obstet Gynecol 2009;34:387-394

10. Rijken MJ, Lee SJ, Boel ME, Papageorghiou AT, Visser GH, Dwell S, Kennedy S, Singhasivanon P, White N, Nosten F, Mcgready R. Obstetric ultrasound scanning by local health workers in a refugee camp on the Thai-Burmese border. Ultrasound Obstet Gynecol 2009;34:395-403

11. Policiano C, Mendes J, Fonseca A, Barros J, Martins D, Reis I, Clode N, Graca L. Impact of maternal weight on the intra-observer and inter-observer reproducibility of fetal ultrasound measurements in the third trimester. Int J Gynecol Obstet 2018;140:53-59

12. Perni SC, Chevernak FA, Kalish RB, Magherini-Rothe S, Predanic M, StreltzoffJ, Skupski DW. Intraobserver and interobserver reproducibility of fetal biometry. Ultrasound Obstet Gynecol 2004;24:654-658

13. Practice Bulletin No 156: Obesity in Pregnancy (Correction). Obstet Gynecol 2016; 128 (6):1450

14. National Center for Health Statistics. Prevalence of obesity among adults and youth: United States 1999-2000 through 2015-2016. http://www.cdc.gov/nchs/data/databriefs/db2019.htm. Accessed October 8, 2019.

15. Tsai PJ, Loichinger M, Zalud I. Obesity and the challenges of ultrasound fetal abnormality diagnosis. Best Pract Res Clin Obstet Gynaecol 2015;29:320-327

16. Heer IM, Kumper C, Vogtle N, Muller-Egloff S, Dugas M, Strauss A. Analysis of factor influencing the ultrasonic fetal estimation. Fetal Diagn Ther 2008; 23:204-210

17. Farrell T, Holmes R, Stone P. The effect of body mass index on three methods of fetal weight estimation. BJOG 2002;109:651-657

18. Cody F, Unterscheider J, Daly S, Geary MP, Kennelly M, McAuliffe F, O'Donoghue K, Hunter A, Morrison J, Burke G, Dicker P, Tully E, Malone FD. The effect of maternal obesity on sonographic fetal weight estimation and perinatal outcome in pregnancies complicated by fetal growth restriction. J Clin Ultrasound 2016; 44 (1):34-9

19. Aksoy H, Aksoy U, Karadag OI, Yucel B, Aydin T, Babayigit MA. Influence of maternal body mass index on sonographic fetal weight estimation prior to scheduled delivery. J Obstet Gynaecol Res 2015;41:1556-1561

20. Kritzer S, Magner K, Warshak CR. Increasing maternal body mass index and the accuracy of sonographic estimation of fetal weight near delivery. J Ultrasound Med 2014;33:2173-2179

21. Fox N S, Bhavsar V, Saltzman DH, Rebarber A & Chasen ST. Influence of Maternal Body Mass Index on the Clinical Estimation of Fetal Weight in Term Pregnancies. Obstetrics & Gynecology 2009; 113(3): 641-645.

22. Paganelli S, Soncini E, Comitini G, Palomba S, La Sala GB. Sonogrpahic fetal weight estimation in normal and overweight / obese healthy term pregnant women by gestation-adjusted projection (GAP) method. Arch Gynecol Obstet 2016; 293(4):775-81

23. Committee Opinion No 688: Management of Suboptimally Dated Pregnancies. Obstet Gynecol 2017; 129 (3):e29-e32

24. Committee Opinion No 700: Methods for Estimating the Due Date. Obstet Gynecol 2017; 129 (5): e150-e154

25. Abuhamad A, Minton K, Benson C, Chudleigh T, Crites L, Doubilet P, Driggers R, Lee W, Mann K, Perez J, Rose N, Simpson L, Tabor A, Benacerraf B. Obstetric and gynecology ultrasound curriculum and competency assessment in residency training programs: consensus report. Am J Obstet Gynecol 2018; 218 (1): 29-67

26. Drukker L, Droste R, Chatelain P, Noble JA, Papageorghiou AT. Expected-value bias in routine third-trimester growth scans. Ultrasound Obstet Gynecol 2020, 55: 375 -382

27. Hadlock FP, Deter RL, Harrist RB, Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. Radiology 1984;152(2):497-501

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

29. Magann EF, Chauhan SP, Bofill JA, Martin JN. Comparability of the amniotic fluid index and single deepest pocket measurements in clinical practice. Aust N Z J Obstet Gynaecol 2003; 43(1):75-7

30. Hughes DS, Magann EF, Whittington JR, Wendel MP, Sandlin AT, Ounpraseuth ST. Accuracy of the ultrasound estimate of the amniotic fluid volume (amniotic fluid index and single deepest pocket) to identify actual low, normal, and high amniotic fluid volumes as determined by quantile regression. J Ultrasound Med 2020; 39(2): 373-378.

31. Sande JA, Ioannou C, Sarris I, Ohuma E, Papageorghiou AT. Reproducibility of measuring amniotic fluid index and single deepest vertical pool throughout gestation. Prenat Diagn 2015; 35(5):434-9

32. Nabhan AF, Abdelmoula YA. Amniotic fluid index versus single deepest vertical pocket as a screening test for preventing adverse outcome. Cochrane Dataase Syst Rev 2008; 16(3): CD006593.

33. Kehl S, Schelkle A, Thomas A, Puhl A, Megdad K, Tuschy B, Berlit S, Weiss C, Bayer C, Heimrich J, Dammer U, Raabe E, Winkler M, Faschingbauer F, Beckmann M, Sutterlin M. Single deepest vertical pocket or amniotic fluid index as evaluation test for predicting adverse pregnancy outcome (SAFE trial): a multicenter, open-label, randomized controlled trial. Ultrasound Obstet Gynecol 2016; 47(6):674-9

34. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Statistics in Medicine 1998; 17:101-110.

35. Stothard K, Tennant T, Bell R & Rankin, J. Maternal Overweight and Obesity and the Risk of Congenital Anomalies. Journal of the American Medical Association 2009; 301(6).

36. Neto MA, Roncato P, Nastri CO, Martins WP. True reproducibility of ultrasound techniques (TRUST): systematic review of reliability studies in obstetrics and gynecology. Ultrasound Obstet Gynecol 2015; 46:14-20

37. Sarris I, Ioannou C, Dighe M, Mitdieri A, Oberto M, Qingqing W, Shah J, Sohoni S, Zidjali W, Hoch L, Altman D, Papageorghiou AT. Standardization of fetal ultrasound biometry measurements: improving the quality and consistency of measurements. Ultrasound Obstet Gynecol 2011; 38: 681-687

38. Yang F, Leung KY, Lee Y, Chan H, Tang M. Fetal biometry by an inexperienced operator using two and three dimensional ultrasound. Ultrasound Obstet Gynecol 2010; 35: 566-571

39. Sarris I, Ohuma E, Ioannou C, Sande J, Altman D, Papageorghiou AT. Fetal biometry: how well can offline measurements from three-dimensional volumes substitute real-time two-dimensional measurements? Ultrasound Obstet Gynecol 2013; 42: 560-570

40. Dashe J, McIntire D, Twickler D. Effect of maternal obesity on the ultrasound detection of anomalous foetuses. Obstet and Gynecol 2009; 113(5):1001-1007

Figures

Tables

Table 1. Maternal age, gestational age and mean BMI according to BMI category.

Characteristic		BMI Category			
	Total (n=110)	18-24.9 (n=20)	25-29.9 (n=30)	30-39.9 (n=30)	≥40 (n=30)
Gestational Age (weeks)	36.7	36.7 ± 1.7	36.4 ± 1.4	36.5 ± 1.6	37.0 ± 1.6
Maternal BMI mean (median)	33.5 (31.0)	23.6 ± 1.1 (23.6)	27.1 ± 1.3 (26.5)	34.2 ± 2.9 (34.4)	45.8 ± 5.0 (45.0)
Age	25.8 ± 5.2	24.0 ± 4.9	25.3 ± 5.7	26.6 ± 4.5	26.9 ± 5.4

Table 2. Descriptive statistics, including interobserver correlation coefficient and reliability coefficients, for interobserver fetal biometric measurements taken by sonographers at EVMS-MFM unit.

Fetal measurement	Mean ± SD	Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
BPD1	8.9 ± 0.42	0.87 (0.820, 0.910)	0.93	0.05 (-1.34 to 1.44)
BPD2	8.9 ± 0.41			
HC1	32.3 ± 1.36	0.72 (0.620, 0.802)	0.84	-0.28 (-2.28 to 1.73)
HC2	32.6 ± 1.40			
AC1	32.8 ± 2.13	0.73 (0.630, 0.810)	0.85	0.04 (-2.25 to 2.34)
AC2	32.7 ± 2.13			
FL1	6.8 ± 0.46	0.74 (0.640, 0.811)	0.85	-0.26 (-2.39 to 1.87)
FL2	6.9 ± 0.42			
EFW1	2894.4 ± 475.3	0.81 (0.740, 0.870)	0.89	-0.09 (-1.57 to 1.39)
EFW2	2925.3 ± 440.5			

Biparietal diameter (BPD), Head circumference (HC), Abdominal circumference (AC), Femur length (FL), Estimated fetal weight (EFW)

[a] The mean z-score difference describes how many standard deviations about the mean of each fetal measurement is the average of the two sonographer measurements.

Table 3. Amniotic fluid index and maximum vertical pocket measurements

Fetal measurement	Mean ± SD	Inter-CC (95%CI)	Reliability Coefficient	Z-score[a]
AFI1	11.6± 4.60	0.65 (0.520, 0.750)	0.78	0.78
AFI2	13.8± 5.30			
MVP1	5.1± 1.50	0.49 (0.340, 0.620)	0.66	0.54
MVP2	5.3± 1.60			

Amniotic fluid index (AFI), Maximum vertical pocket (MVP)

[a] The z-score describes how many standard deviations about the mean of each fetal measurement is the average of the two sonographer measurements.

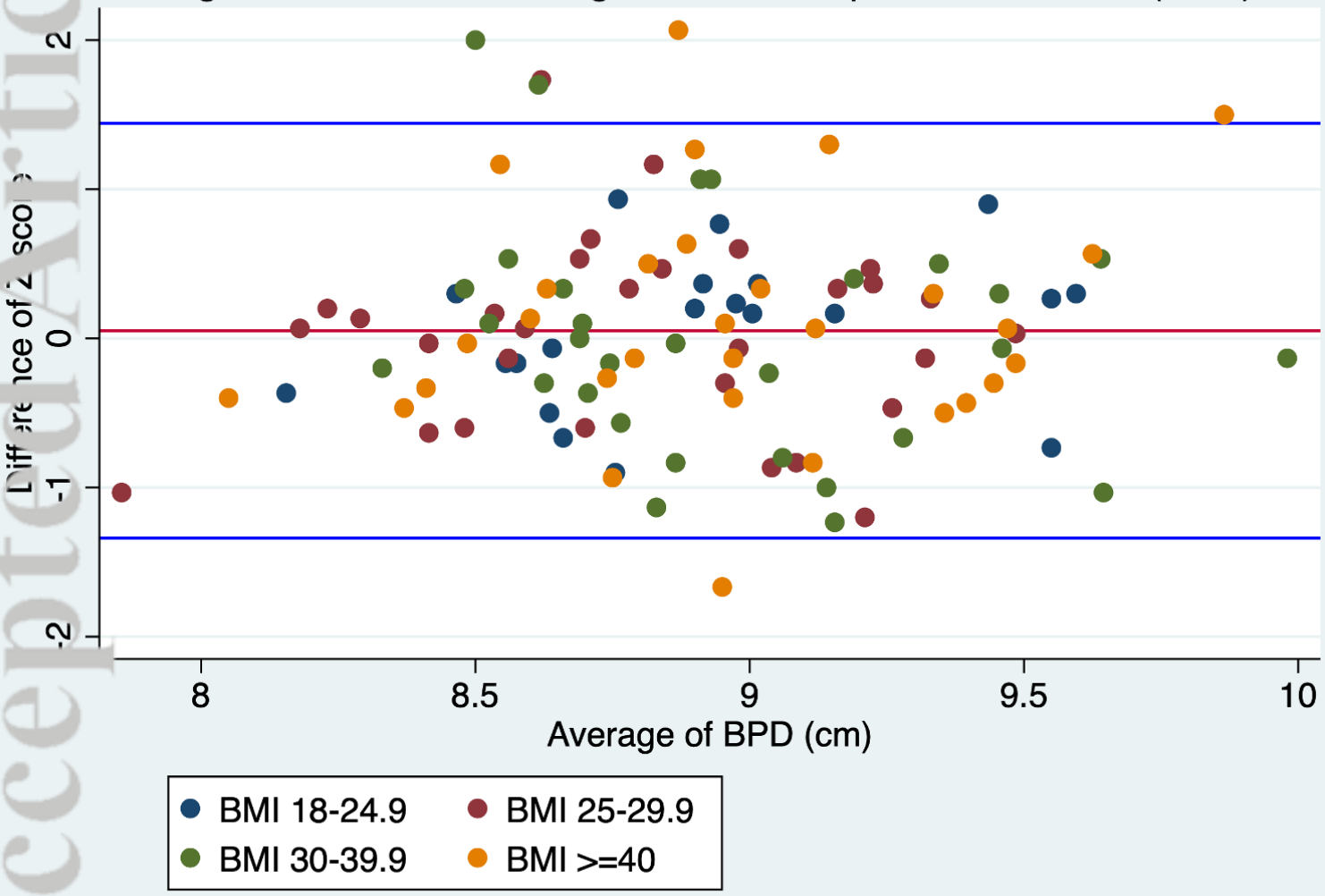
Table 4. The interobserver correlation coefficient and the corresponding reliability coefficients for the fetal biometric measurements by BMI category, expressed as Z-scores.

Parameter	BMI Category	Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
BPD	18-24.9	0.93 (0.830, 0.970)	0.96	0.07 (-0.96 to 1.10)
	25-29.9	0.89 (0.780, 0.950)	0.94	0.02 (-1.25 to 1.30)
	30-39.9	0.85 (0.710, 0.930)	0.91	0.01 (-1.56 to 1.57)
	> 40	0.85 (0.710, 0.930)	0.92	0.11 (-.018 to 0.40)
Parameter		Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
HC	18-24.9	0.59 (0.200, 0.820)	0.74	-0.42 (-2.21 to 1.37)
	25-29.9	0.72 (0.480, 0.860)	0.84	-0.33 (-2.47 to 1.80)
	30-39.9	0.75 (0.540, 0.870)	0.86	-0.23 (-2.10 to 1.63)
	> 40	0.71 (0.480, 0.850)	0.83	-0.17 (-2.38 to 2.05)
Parameter		Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
AC	18-24.9	0.44 (-0.010, 0.740)	0.60	-0.11 (-2.96 to 2.73)
	25-29.9	0.57 (0.260, 0.770)	0.72	0.04 (-2.66 to 2.73)
	30-39.9	0.83 (0.670, 0.920)	0.91	0.14 (-1.77 to 2.04)
	> 40	0.86 (0.720, 0.930)	0.92	0.06 (-1.81 to 1.92)
Parameter		Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
FL	18-24.9	0.65 (0.029, 0.850)	0.78	-0.63 (-3.36 to 2.10)
	25-29.9	0.79 (0.590, 0.890)	0.87	-0.40 (-0.67 to -0.14)
	30-39.9	0.77 (0.570, 0.880)	0.87	-0.03 (-2.40 to 2.34)
	> 40	0.71 (0.470, 0.850)	0.82	-0.08 (-2.02 to 1.85)
Parameter		Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
AFI	18-24.9	0.65 (0.300, 0.850)	0.79	N/A
	25-29.9	0.76 (0.550, 0.880)	0.86	N/A
	30-39.9	0.49 (0.150, 0.720)	0.66	N/A
	> 40	0.61 (0.320, 0.800)	0.74	N/A
Parameter		Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
MVP	18-24.9	0.73 (0.420, 0.880)	0.83	N/A
	25-29.9	0.54 (0.230, 0.760)	0.69	N/A
	30-39.9	0.32 (-0.050, 0.620)	0.50	N/A
	> 40	0.33 (-0.004, 0.620)	0.49	N/A
Parameter		Inter-CC (95%CI)	Reliability Coefficient	Mean Z-score Difference[a]
EFW	18-24.9	0.68 (0.340, 0.860)	0.81	-0.30 (-1.94 to 1.34)
	25-29.9	0.69 (0.440, 0.840)	0.82	-0.12 (-1.79 to 1.56)
	30-39.9	0.90 (0.800, 0.950)	0.95	-0.01 (-1.18 to 1.56)
	> 40	0.84 (0.680, 0.920)	0.90	0.01 (-1.45 to 1.47)

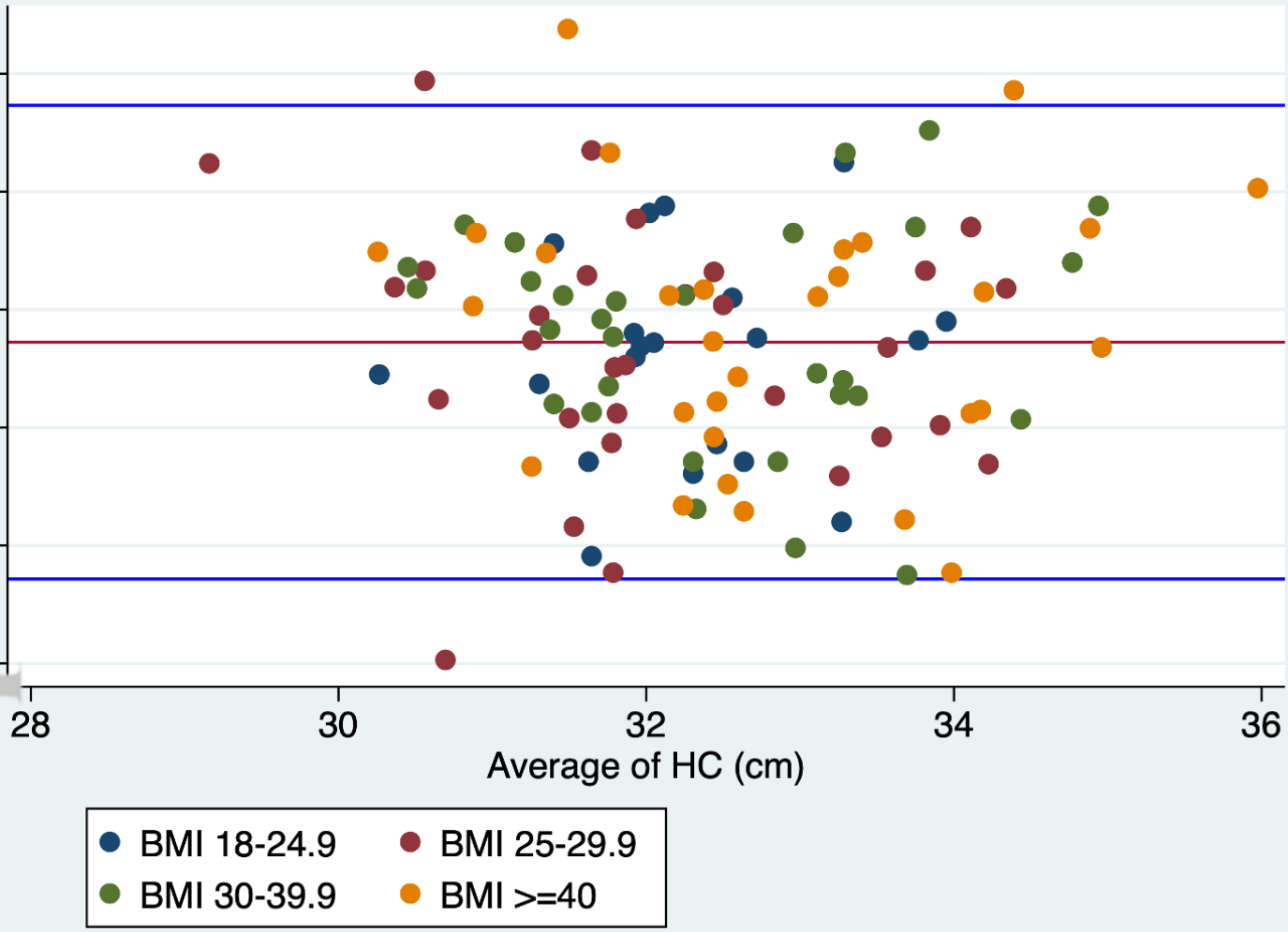
Body Mass Index (BMI), Biparietal diameter(BPD), Head circumference (HC), Abdominal circumference (AC), Femur length (FL), Amniotic fluid index (AFI), Maximum vertical pocket (MVP), Estimated fetal weight (EFW)

[a] The mean z-score difference describes how many standard deviations about the mean of each fetal measurement is the average of the two sonographer measurements.

Figure 1. Interobserver agreement for biparietal diameter (BPD)

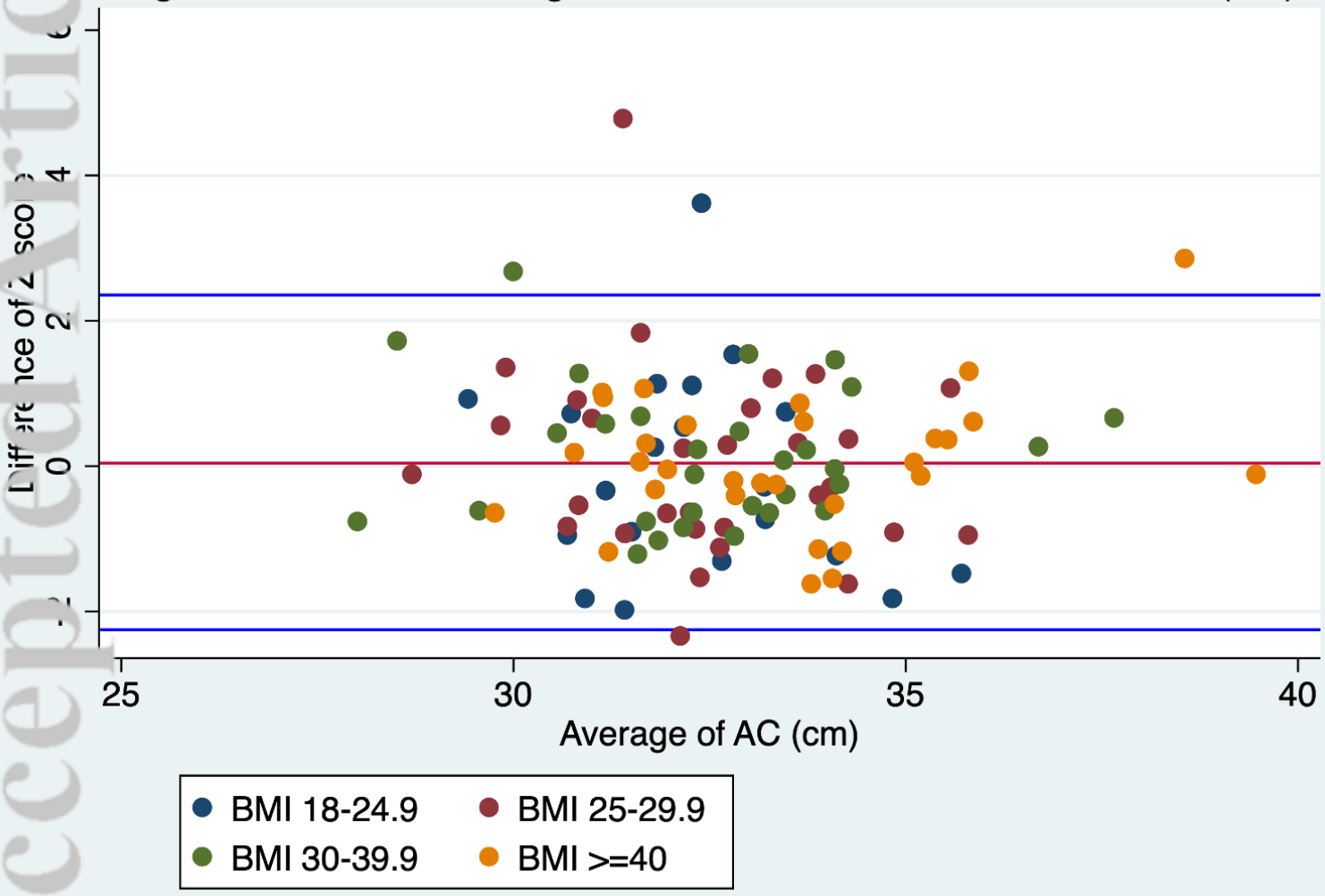


UOG_23646_Figure1-400dpi.png



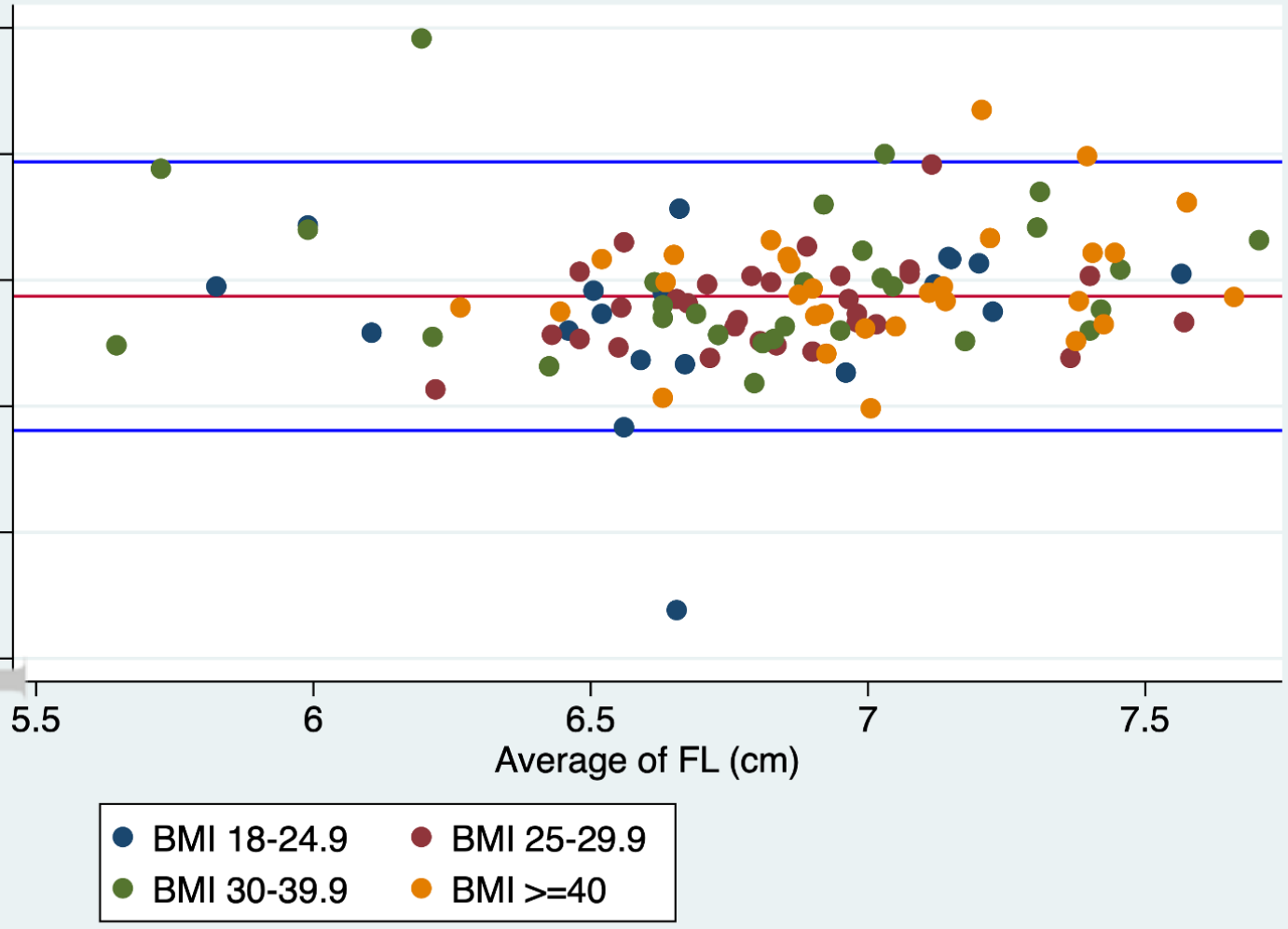
UOG_23646_Figure 2-400dpi.png

Figure 3. Interobserver agreement for abdominal circumference (AC)



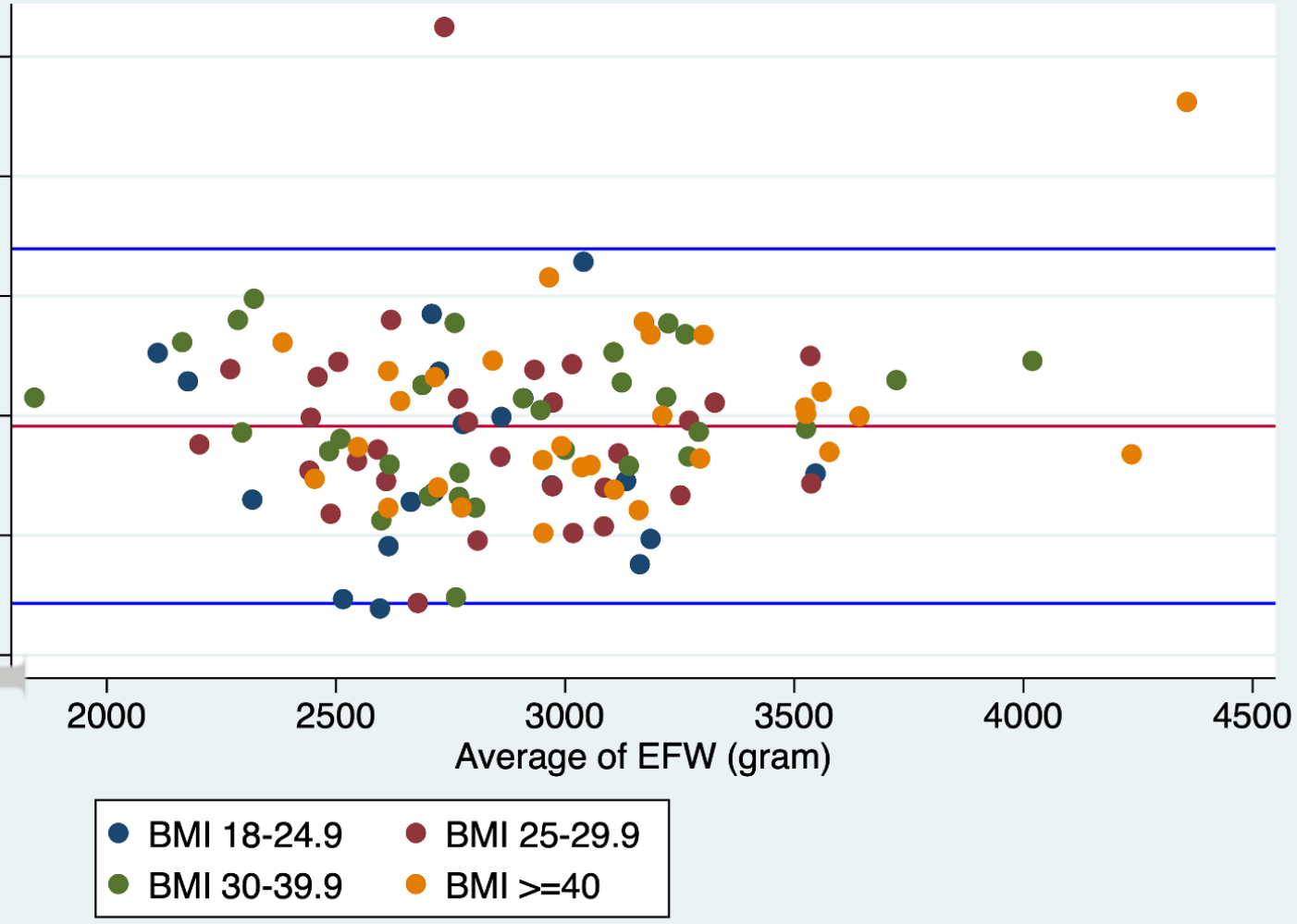
UOG_23646_Figure 3-400dpi.png

Figure 4. Interobserver agreement for femur length (FL)



UOG_23646_Figure 4-400dpi.png

Figure 5. Interobserver agreement for estimated fetal weight (EFW)



UOG_23646_Figure 5-400dpi.png