

peer reviewed ORIGINAL ARTICLE

Inter-operator reproducibility of ultrasound measurement of three foetal biometric parameters: comparison of experienced sonographers and radiography interns

NO Egbe¹ Ph.D | EU Ekpo¹ B.Sc | CC Nzotta² Ph.D¹Department of Radiography & Radiological Science, University of Calabar, Calabar, Nigeria.²Department of Radiography/Radiological Sciences, Nnamdi Azikiwe University, Awka, Nigeria

Abstract

Background: The dearth of experienced sonographers in Nigeria has created an all-comers environment in which many would-be sonographers enter the field with minimal supervision giving rise to the question of competence of personnel and, often, the accuracy of results obtained.

Objective: To assess inter-operator reproducibility in the measurement by sonographers and radiography interns of three common foetal biometric parameters: bi-parietal diameter (BPD), femur length (FL), and abdominal circumference (AC).

Method: Twenty (n=20) women, carrying normal singleton fetuses of between 20 and 40 weeks gestational age, consented to participate in the study. They were evaluated in three ultrasound centers. Ultrasound scans were performed on all of them. Biometric measurements were obtained following standard protocols. Foetal biometric measurements (bi-parietal diameter (BPD), femur length (FL), and abdominal circumference (AC) were performed independently by one sonographer and one intern who were blinded to the purpose of the study and the figures of the other participant in each centre. Coefficient of variability percent was used to assess inter-operator variability. Intra-class correlation coefficient (ICC) at a 95% confidence interval was used as a metric to compute weighted kappa (Kw) to assess inter-operator reproducibility of measurements.

Results: Coefficient of variability (COV %) ranged from 0.84 – 1.79, and the mean coefficient of variability was below 2.0 % in all locations for all measured parameters. Intra-class correlation coefficient, indicating between operator agreement ranged from 0.95 – 0.98 for BPD, 0.90 – 0.91 for FL, and 0.92 – 0.99 for AC. The overall agreement for all centers was 0.94 ± 0.04 ; $P=0.0001$.

Conclusion: Sonographic measurements of fetal biometric parameters measures of experienced sonographers are highly reproducible by interns undertaking ultrasound training.

Keywords

Gestational age, obstetric ultrasound, femur length, bi-parietal diameter, abdominal circumference

Introduction

Technological innovations in medical imaging have enabled evaluation of maternal reproductive health and foetal developmental changes that occur *in-utero*.^[1] Ultrasound in particular is an exceptionally important imaging tool in the field of obstetrics and gynaecology for detection and characterisation of pelvic abnormalities.^[2] It has now become the cornerstone of foetal and maternal health assessment.^[3,4] Increased availability, portability and relative low cost, have led to the widespread use of ultrasound for monitoring gestation from the early to late stages and, for making clinical decisions about care during labour. Ultrasound has been shown to be the most reliable imaging tool for assessment of foetal viability, multiple gestations, and for detection of both intra and extra-uterine conditions that may constitute risk to the developing foetus.^[5-7] Importantly, foetal biometric parameters,

for example, crown-rump length, femur length, and head and abdominal measurements, using ultrasound provide information about foetal growth and gestational age (GA).^[4,8]

Foetal growth, defined as the time dependent changes in body dimensions that occur throughout the gestational period, occurs more rapidly in the first and second trimester.^[8,9] It is increasingly important to evaluate such changes as they occur to assess whether such foetal development changes are concomitant with gestational age. Ultrasound foetal biometry (the measurement of the growth and development of parts of foetal anatomy) provides information about GA. It is also invaluable in detection of growth aberration such as intra-uterine growth retardation and anatomical deformities in the later stages of pregnancy.^[3,7] Although GA can be estimated from the last menstrual period (LMP), the accuracy of such estimates may be limited by a patient's memory deficit.

Clinically, GA is widely estimated using fundal height,^[10] and maternal parameters such as uterine fibroid and polyhydramnios, which may influence estimates and cause inaccuracies in GA estimation.^[6] These deficiencies inspired the use of ultrasound for foetal biometry. Ultrasound is safe and is the most reliable and widespread tool used for estimation of GA, especially in women who do not remember the date of their LMP or whose fundal height on abdominal examination does not correspond to dates.^[11] In addition to providing information about foetal parameters, ultrasound is also useful for evaluation of the placenta and cervix during pregnancy.^[12]

Foetal biometry is helpful in accurately predicting gestational age during the first trimester.^[13,14] A number of foetal biometric parameters can be used to estimate gestational age and include crown-rump-length, femur length, foetal head parameters such as bi-parietal diameter and head

circumference, and abdominal parameters such as abdominal circumference and transverse abdominal diameter.^[15,16] It is important that GA, which is estimated on the same day, should be the same regardless of the ultrasound equipment used or the medical personnel performing ultrasound GA assessment. However, accurate GA estimation using ultrasound requires accurate measurement of these parameters. In other words, a reliable prediction of foetal growth and GA by ultrasound depends on the competence of an operator. Variation in measurement procedures could lead to inaccuracy in GA estimation and/or inter-operator differences in foetal dating, which may cause differences in clinical decision-making, including early inducement of delivery, wrong choices of delivery approaches^[17] and even foetal death resulting from prolonged intra-uterine stay.^[18] These consequences of wrong GA estimation underscore the need for accurate and reproducible foetal biometry using ultrasound.

Despite a long history of utilising ultrasound to estimate GA, and a number of studies covering several foetal anatomy, physiology and related issues,^[16,19,20] there are no studies documenting inter-operator reproducibility of biometric parameters in the Nigerian environment. Also, the dearth of radiologists and qualified sonographers has become a motivation for many young graduate radiographers to undertake clinical ultrasound duties, sometimes without adequate training and supervision. Whilst it is desirable to have personnel to perform ultrasound examination, it is more relevant that such personnel produce acceptable and reproducible performance outcomes.

However, the performance of these graduate radiographers in foetal ultrasonography is unknown thus emphasises the need for research. This study sought to determine the inter-operator reproducibility of ultrasound measurement of foetal biometric parameters using standard measure-

ment landmarks. In particular, the study sought to assess how well intern radiographers are able to reproduce foetal biometric measurements of trained sonographers.

Materials and methods

The study was approved by the institution (IRB: NAU 2015/0278). It involved expectant women undergoing routine ultrasound examinations. Prior to the commencement of the study, the procedure was explained to all the women. Informed consent was obtained from those who agreed to participate in the study. The study was performed in three ultrasound scan centers. The respective ultrasound equipment used at the three centers produced similar foetal age and estimated delivery date (EDD) for the same measurement value of each biometric parameter. In other words, all equipment outputted the same GA and EDD for the same measurement value of femur length. QA testing for distance measurement was made by comparing the GA outputted for each measured parameter to the established reference obstetric chart.^[21] This was to ensure that the equipment produced the correct GA for a given value of the obstetric parameter measurement.

Participants' characteristics

The participants were women who presented for routine obstetric ultrasound scans, and who had consented to participate in the study. The criteria for selection of expectant women were: single foetus in-utero; between 20-40 weeks pregnant as calculated from LMP at the time of scan; and healthy with no previous history of foetal congenital anomalies. Women having foetuses with suspicious anomalies were excluded as were those who could not provide details of their clinical history. Of the women that volunteered to participate in the study, only 20 women fulfilled the inclusion criteria and were used for the study. To assess the effect of sample size, power analysis was performed using alpha metric at 95% confidence interval.

The 20 participants attended scanning sections in the three ultrasound centers one day apart.

The imaging personnel were one experienced sonographer and an intern radiographer in each center (study site). The interns had a fair knowledge in ultrasonography following a 2-3 month period of intensive training. The sonographers were in possession of postgraduate certificates in sonography, with an average experience in sonography of 5.5 ± 2.1 years. The interns, on the other hand, were all holders of a bachelor's degree in radiography, with at least two months' intensive exposure to ultrasonography. Two of the interns were five months post-graduation, and the third had just graduated. All sonographers and interns were blinded to their patients' LMP.

Scanning procedure

Ultrasound scanners used were dedicated real-time systems and were checked to ensure they produced a consistent output of biometric measurements. The ultrasound equipment and probes used, as well as operator sonographic experience, are shown in Table 1. All participants underwent trans-abdominal ultrasound evaluation in each of the centres. Operator measurements for bi-parietal diameter (BPD), femur length (FL) and abdominal circumference (AC),^[22] recorded in millimetres (mm) were performed on them following the standard protocol as detailed in the British Medical Ultrasound Society (BMUS) guidelines.^[23] These three foetal biometric parameters were used in the current study because they are the routinely used parameters for foetal dating in this region. Reasons for their use include that they are easy to determine and are used by the ultrasound system to compute foetal weight.

The BPD was measured by obtaining an oval image of each foetal skull depicting undistorted falx cerebri, thalami and the cavum septum pellucidum. Measurement was done by extending the measuring cal-

Table 1: Scanning equipment and operator experience in the three centres.

CENTRE CODE	US EQUIPMENT	SONOGRAPHER EXPERIENCE	INTERN EXPERIENCE	PROBE TYPE
Centre 1	Medison 600A	7 years	5 months	3.5 MHz
Centre 2	Mindray DP 3300	5 years	FG	3.5 MHz
Centre 3	Shimadzu 600	6 years	5 months	3.5 MHz

FG – Fresh graduate

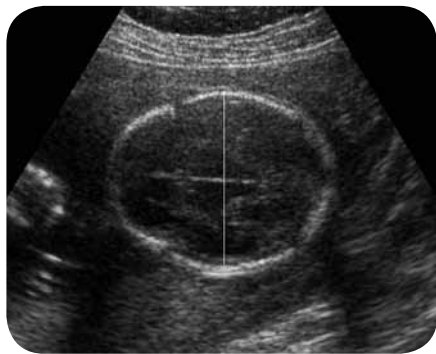


Figure 1. Method of measurement of biparietal diameter (BPD).



Figure 2. Measurement of femoral length (FL).

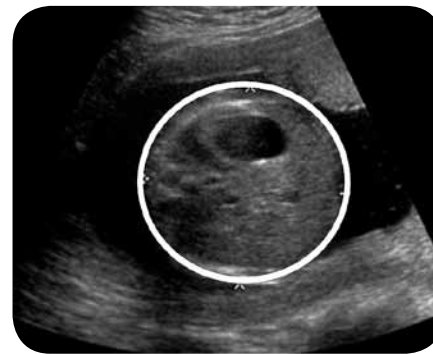


Figure 3. Measurement of abdominal circumference (AC).

liper from the outer table of the foetal skull above, to the opposite inner table, at the level of the thalami (Figure 1).

The FL was measured as shown in Figure 2, by obtaining an image of the femur showing its proximal and distal ends (round, echogenic cartilaginous femoral head and femoral condyles). The measuring callipers were extended from the femoral head to the lateral femoral condyle (straight aspect of the femur).

The AC was measured by obtaining a circular/transverse image of the foetal abdomen showing the short umbilical vein and the stomach. An elliptical calliper measurement method (Figure 3) was used to avoid the spuriousness associated with the trace method.

Individual operators, in each centre, scanned all 20 participants. They recorded their respective foetal biometric data and made hard copies. Scans were performed independently by each imaging operator and in-turn. On screen measured values were deleted before switching operators. None of the operators were not allowed access to the other operators' results. The scan time of each operator was recorded.

Statistical analysis

The Statistical Package for Social Sciences (SPSS) Version 18 was used for data analysis. For each measurement made, the coefficient of variability percent (COV%) was used to determine the variance between operators. Mean COV% was determined per parameter in each location. From the sets of each operator's measurements, the inter-operator reproducibility was determined by single measure intra-class correlation coefficient (ICC) as a metric for weighted kappa (K_w). Bootstrapping was used to compute the 95% confidence interval for all measures. Statistical dif-

ference between operators in each location was assessed based on a two-sided p-value ≤ 0.05 .

In calculating the level of reproducibility, the intra-class correlation (ICC) was used a metric for the degree of weighted agreement (K_w) between pairs of observations by the operators as described by Landis and Koch.^[24] Intra-class reliability coefficient values were interpreted as slight (0.00-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80) and perfect (0.81-1.00).

Results

In total, sixty ($n=60$) obstetric scans for foetal biometry were performed (twenty in each of the three study sites/locations). Power analysis demonstrated an observed power of 78%, indicating that the sample size can be relied upon to generalise findings of the study. The average scanning time per participant across the three sites was 18 ± 4 minutes for the sonographers and 24 ± 5 minutes for the interns.

The mean values for measurements of BPD, FL and AC by the sonographers and interns are presented in Figures 4-6, respectively. Careful inspection shows that the measurements by the operators were fairly evenly matched for each participant. The mean percentage coefficient of variability (COV%) between operators was

1.22 for BPD, with the respective location values, 1.27, 0.84 and 1.55 for locations 1, 2 and 3, respectively. Measurement of FL averaged for all operators across the three locations varied by 1.38%, with the respective location values of 1.47% (location 1), 1.22% (location 2), and 1.46% (location 3). Abdominal circumference (AC) measurement demonstrated a COV% of 1.50, with respective location values of 1.79, 1.58 and 1.14 for locations 1, 2 and 3, respectively.

Table 2 shows values obtained for intra-class correlation coefficients (ICC). The results demonstrate high inter-operator consistency and reproducibility in the measurements of foetal biometric parameters: BPD, FL and AC. Intra-class correlation coefficient, indicating inter-operator agreement ranged from 0.95 – 0.98 for BPD, 0.90 – 0.91 for FL, and 0.92 – 0.99 for AC. The p-values representing the levels of statistical significance for these comparisons are shown in Table 2. The overall agreement for all sites was 0.94 ± 0.04 ; $p=0.0001$. There was no statistically significant difference in biometric measurements of each respective sonographer and intern in the three locations.

Discussion

The expected delivery date of a pregnancy is often determined from the GA. Different clinical decisions, concerning deliv-

Table 2: Single measure Intra-class correlation coefficient (ICC).

PARAMETER	LOCATION MEASUREMENT		
	LOCATION 1	LOCATION 2	LOCATION 3
BPD	0.95	0.98	0.95
FL	0.90	0.91	0.91
AC	0.92	0.99	0.99

ery time and modes, are made based on the estimated GA, where women deemed to be post-dated are induced or undergo caesarean sections.^[10] Therefore, it is increasingly important that sonographic estimates of GA are accurate and consistent to reduce variability from ultrasound biometric measures and improve use of such data in clinical decision-making. In consideration of this, the current study examined the reproducibility of ultrasound foetal biometric parameters measurement by comparing data of trained sonographers and intern radiographers. The results demonstrate high inter-operator reproducibility of foetal biometric measurement. They also show that the measurements of experienced sonographers appear to be well matched by those of the inexperienced operators undertaking training in ultrasound (Figures 4-6).

Studies elsewhere have shown high reproducibility of foetal biometric parameters. One study assessed intra- and inter-observer reproducibility of early foetal growth parameters (CRL and FL) in 21 singleton pregnancies and demonstrated high intra- and inter-observer correlation.^[25] In another study, Babieri et al^[26] reported a good level of inter-examiner agreement measurement of the cross-sectional diameter and area of the umbilical cord and its vessels in 221 pregnant women. A recent study^[27] however suggests that good agreement may be dependent on the mode of expression of measurements. Therefore, simple parameters chosen for the current study may have made it easier for the interns and could be responsible for the high inter-operator agreement.

Since ultrasonography is operator dependant, it is important to maintain a low level of variability in foetal dating. Factors such as training, mentoring, and experience have been shown to impact performance in radiology.^[28,29] In particular, a recent study shows the impact of mentorship on performance for radiographers in the private sector.^[29] Evidence from the respective locations/sites in the current study revealed that the interns had each spent numerous hours of mentor-guided training in a private ultrasound setting. This level of commitment may have accounted for their high level of agreement with experienced operators (Table 2). Also, the fact that the study immediately followed their training may have also positively impacted the

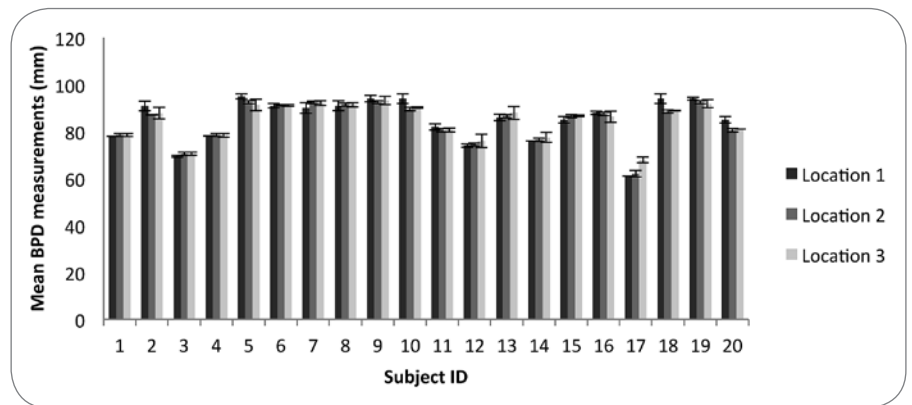


Figure 4. Mean values of BPD measurements of the two operators in each location. Error bars are 1 standard deviation.

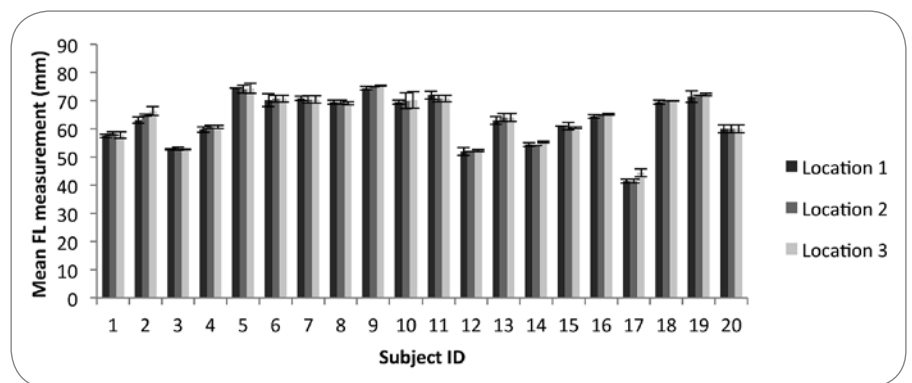


Figure 5. Mean values of FL measurements of the two operators in each location. Error bars are 1 standard deviation.

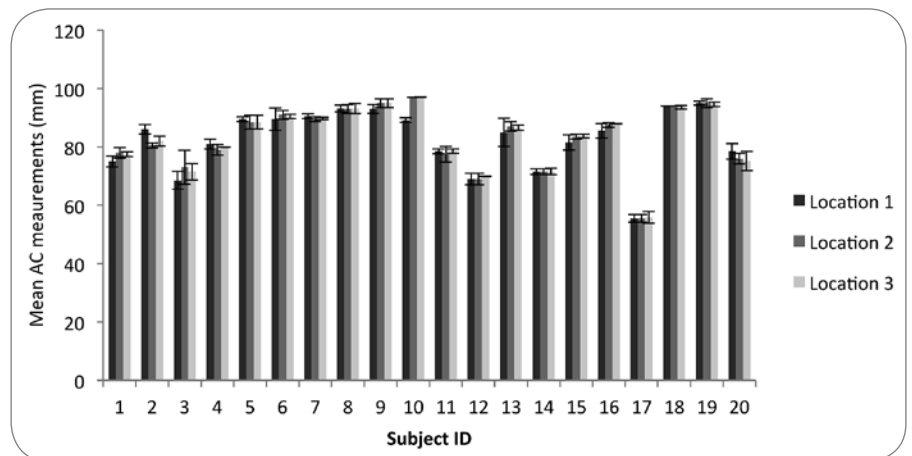


Figure 6. Mean values of AC measurements of the two operators in each location. Error bars are 1 standard deviation.

performance of the interns, and suggests that with adequate training, interns can deliver reliable results in the ultrasound foetal biometric studies, at least for BPD, FL and AC.

Whilst the findings of the current study are encouraging, it should be remembered that they are specific to foetal obstetric parameter measurement and do not demonstrate that radiography interns have the skills to produce and interpret ultrasound

images. Since ultrasonography is not limited to the three parameters of biometric measurements used in the current study, a mere two month exposure to some components may not guarantee the acquisition of sufficient competency in the modality. We hypothesise that if training and adequate mentorship is made available to undergraduate radiography students, with emphasis on modality specialisation, that they would acquire the basic skills required for ultrasound examinations.

Currently, radiography training is a five year programme with the recent introduction of ultrasonography as a standalone course. The ultrasound course is intended to provide theoretical and practice-based training in ultrasound image acquisition and interpretation. It is believed that this development, pursued with commitment to teaching and clinical exposure, should ensure that student radiographers acquire adequate ultrasonography competency at graduation. Whilst this does not rule out a further postgraduate training and certification, early exposure would provide a good foundation for postgraduate training in sonography. The impact of the on-going training and mentorship interventions on students' clinical competency in the interpretation of ultrasound images should be explored in future studies.

Conclusion

The results demonstrate high inter-operator consistency in measurement of three foetal biometric parameters widely used for gestational age assessment between trained sonographers and interns undertaking ultrasound training. Data produced suggest that the sonography manpower needs of the nation may stand to benefit from postgraduate ultrasound training of radiography interns.

Acknowledgement

The authors are grateful to all the participants for agreeing to participate in this study; the management and staff of the three sonography sites and the participating sonographers and interns.

References

- Merton DA. Diagnostic medical ultrasound technology: a brief historical review. *J Diag Med Sonography* 1997;13(5):suppl 10S-23S.
- Guerriero S, Saba L, Ajossa S, Peddes C, Angiolucci M, Perniciano M, et al. Three-dimensional ultrasonography in the diagnosis of deep endometriosis. *Human reproduction (Oxford, England)* 2014; Jun;29(6):1189-98. PubMed PMID: 24664128. Epub 2014/03/26. eng.
- Switzer PJ, James CA, Frettag MA. Value and limitations of obstetrical ultrasound: uncovering abnormalities at earlier stages. *Canadian Family Physician Medecin de famille Canadien* 1992; Jan;38:121-8. PubMed PMID: 21229124. Pubmed Central PMCID: PMC2145611. Epub 1992/01/01. eng.
- Seeds JW. The routine or screening obstetrical ultrasound examination. *Clin Obstet Gynecol* 1996;39(4):814-30. PubMed PMID: 8934033. Epub 1996/12/01. eng.
- Bergen NE, Bouwland-Both MI, Steegers-Theunissen RP, Hofman A, Russcher H, Lindemans J, et al. Early pregnancy maternal and fetal angiogenic factors and fetal and childhood growth: the generation R study. *Human Reproduction (Oxford, England)* 2015; Jun;30(6):1302-13. PubMed PMID: 25854264. Epub 2015/04/10. eng.
- Magann EF, Sandlin AT, Ounpraseuth ST. Amniotic fluid and the clinical relevance of the sonographically estimated amniotic fluid volume: Oligohydramnios. *J Ultrasound Med.* 2011; 30(11):1573-85. PubMed PMID: 22039031. Epub 2011/11/01. eng.
- Yeo L, Guzman ER, Shen-Schwarz S, Walters C, Vintzileos AM. Value of a complete sonographic survey in detecting fetal abnormalities: correlation with perinatal autopsy. *J Ultrasound Med* 2002; 21(5):501-10. PubMed PMID: 12008812. Epub 2002/05/15. eng.
- Simon NV, Levisky JS, Shearer DM, O'Lear MS, Flood JT. Influence of fetal growth patterns on sonographic estimation of fetal weight. *JCU* 1987; 15(6):376-83. PubMed PMID: 3134435. Epub 1987/07/01. eng.
- Khalid S MA, Shahida Z. Foetal biometry. *Pakistan J of Medicine* 2006;22(4):503-8.
- Ananth CV. Menstrual versus clinical estimate of gestational age dating in the United States: temporal trends and variability in indices of perinatal outcomes. *Paediatric and Perinatal Epidemiology* 2007; Sep;21 Suppl 2:22-30. PubMed PMID: 17803615. Epub 2007/09/22. eng.
- Lubusky M PM, Langova M, Vomackova K, Cizek L. Discrepancy of ultrasound biometric parameters of head circumference, biparietal diameter in breech presented fetuses. *Ultrasound in Obstetrics & Gynaecology* 2007;151(2): 323-6.
- Pandipati S, Combs CA, Fishman A, Lee SY, Mallory K, Ianovich F. Prospective evaluation of a protocol for using transabdominal ultrasound to screen for short cervix. *Am J of Obst & Gynecology* 2015; Apr 30. PubMed PMID: 25935775. Epub 2015/05/04. Eng.
- Kimberly B FN, Ken L, Vancouver BC. Determination of gestational age by ultrasound. *J Obstet Gynaecol Can* 2014;36(2):171-81.
- Caughey AB, Nicholson JM, Washington AE. First- vs second-trimester ultrasound: the effect on pregnancy dating and perinatal outcomes. *Am J of Obsts & Gynecology* 2008; 198(6):703 e1-5; discussion e5-6. PubMed PMID: 18538160. Pubmed Central PMCID: PMC2603611. Epub 2008/06/10. eng.
- Fatalah HA Al, Abbas HY, Khafaji MA, Al-safi KGH, Jastaniah SD. Accuracy of ultrasound to determine gestational age in the third trimester. *Open J of Medical Imaging.* 2014; 4:126-32.
- Mador ES PI, Ekedigwe JE, Ogunranti JO. Ultrasonic biometry of Nigerian fetuses I: femur length. *Asian J of Medical Sciences* 2012;4(2):94-8.
- Galal M, Symonds I, Murray H, Petraglia F, Smith R. Postterm pregnancy. Facts, views & vision in *ObGyn.* 2012 ;4(3):175-87. PubMed PMID: 24753906. Pubmed Central PMCID: PMC3991404. Epub 2012/01/01. eng.
- Kallan JE. Effects of interpregnancy intervals on preterm birth, intrauterine growth retardation, and fetal loss. *Social Biology* 1992; 39(3-4):231-45. PubMed PMID: 1340042. Epub 1992/01/01. eng.
- Mador E S ECC, Mutihir J T, Adoga G I, Ogunranti J O. Ultrasonic biometry: biparietal diameter in Nigerian fetuses. *Niger Med Journal* 2011;52(4):258.
- Mador ES PI, Ekedigwe JE, Ogunranti JO. Ultrasonic biometry of Nigerian Fetuses I: estimated fetal weight. *Asian J of Medical Sciences* 2012;4(2):89-93.
- Pam Loughna LC, Tony Evans, Trish Chudleigh. Fetal size and dating: charts recommended for clinical obstetric practice. *Ultrasound* 2009;17(3):161-7.
- Valentine L BI. Intra-observer and inter-observer reproducibility of ultrasound measurement of cervical length and width in the second and third trimesters of pregnancy. *Ultrasound Obstet Gynaecol* 2002;20(3):256-62.
- G H. The safe use of ultrasound in medical diagnosis 3rd Edition UK. *British Institute of Radiology*; 2012.
- Landis JR KG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74
- Verbog BO MP, Hofman A, Jaddoe VVV, Witteman JCM, Steegers EAP. Intra-observer and inter-observer reproducibility study of early foetal growth parameters. *Prenatal Diagnosis* 2008;28(4):323-31.
- Barbieri C CJ, Souza CE, Marussi EF and Costa JV. Inter- and intra-observer variability in sonographic measurements of the cross-sectional diameters and area of the umbilical cord and its vessels during pregnancy. *Reproductive Health* 5:5.
- Sarris I IC, Chamberlain P, Ohima E, Roseman F, Hach L, Altman DG, PapaGeorghiou AT. Intra and inter observer variability in fetal ultrasound measurements. *Ultrasound Obst Gynaecol* 2012;39 266-73.
- Nodine CF, Kundel HL, Mello-Thoms C, Weinstein SP, Orel SG, Sullivan DC, et al. How experience and training influence mammography expertise. *Acad Radiol* 1999; 6(10):575-85. PubMed PMID: 10516859. Epub 1999/10/12. eng.
- Ekpo EU, Egbe NO, Akpan BE. Radiographers' performance in chest X-ray interpretation: the Nigerian experience. *BJR* 2015; June 4:20150023. PubMed PMID: 25966290. Epub 2015/05/13. Eng.