



THE AGA KHAN UNIVERSITY

eCommons@AKU

Imaging & Diagnostic Radiology, East Africa

Medical College, East Africa

May 2015

Reproducibility of measuring amniotic fluid index and single deepest vertical pool throughout gestation

Joyce Sande

Aga Khan University, joyce.sande@aku.edu

C. Ioannou

University of Oxford

I. Sarris

University of Oxford

E. O. Ohuma

University of Oxford

A. T. Papageorghiou

University of Oxford

Follow this and additional works at: http://ecommons.aku.edu/eastafrica_fhs_mc_imaging_diagn_radiol

 Part of the [Radiology Commons](#)

Recommended Citation

Sande, J., Ioannou, C., Sarris, I., Ohuma, E. O., Papageorghiou, A. T. (2015). Reproducibility of measuring amniotic fluid index and single deepest vertical pool throughout gestation. *Prenatal Diagnosis*, 35(5), 434-439.

Available at: http://ecommons.aku.edu/eastafrica_fhs_mc_imaging_diagn_radiol/8

ORIGINAL ARTICLE

Reproducibility of measuring amniotic fluid index and single deepest vertical pool throughout gestation

J. A. Sande^{1,2}, C. Ioannou², I. Sarris², E. O. Ohuma^{2,3} and A. T. Papageorghiou^{2*}

¹Department of Radiology, Aga Khan University Hospital, Nairobi, Kenya

²Oxford Maternal & Perinatal Health Institute (OMPHI), Green Templeton College and Nuffield Department of Obstetrics and Gynecology, John Radcliffe Hospital, University of Oxford, Oxford, UK

³Centre for Statistics in Medicine, University of Oxford, Botnar Research Center, Oxford, UK

*Correspondence to: A. T. Papageorghiou. E-mail: aris.papageorghiou@obs-gyn.ox.ac.uk

ABSTRACT

Objective The aim of this study is to assess the intraobserver and interobserver reproducibility of measurement of amniotic fluid index (AFI) and single deepest vertical pool (SDVP), also known as the maximal vertical pocket.

Methods A total of 175 fetuses were evaluated. For each fetus, two observers acquired duplicate sets of AFI and SDVP. Measurement differences were expressed as actual and percentage values. For all comparisons, Bland–Altman plots were used to compare differences, and limits of agreement were calculated.

Results Intraobserver and interobserver agreement remained fairly constant with gestation, both for AFI and SDVP. The intraobserver limits of agreement for AFI were -5.2 to 5 cm or -39% to 37% ; whereas for SDVP, these were -2.6 to 2.4 cm or -52% to 48% . The interobserver limits of agreement for AFI measurement were -7.3 to 7.1 cm or -54% to 53% and for SDVP measurement were -2.5 to 2.5 cm or -51% to 52% . Intraobserver coefficient of variation for SDVP was 14% and for AFI was 19% ; the interobserver coefficient was 19% for both AFI and SDVP.

Conclusion Limits of agreement for both methods are wide. The choice of method should be dictated by clinical considerations other than method reproducibility. © 2014 John Wiley & Sons, Ltd.

Funding sources: ATP and CI are supported by the Oxford Partnership Comprehensive Biomedical Research Centre with funding from the Department of Health NIHR Biomedical Research Centers funding scheme. This project was supported by the INTERGROWTH-21st grant ID no. 49038 from the Bill & Melinda Gates Foundation to the University of Oxford, for which we are very grateful. Aris Papageorghiou and Christos Ioannou are supported by the Oxford Partnership Comprehensive Biomedical Research Centre with funding from the Department of Health NIHR Biomedical Research Centers funding scheme. The INTERGROWTH-21st study is supported by a grant from the Bill & Melinda Gates Foundation.

Conflicts of interest: None declared

INTRODUCTION

Amniotic fluid is a key indicator of fetal well-being in the second half of pregnancy. It is a function of both urine production and swallowing. At the beginning of pregnancy, amniotic fluid volume (AFV) is higher than fetal volume; the two volumes become equal soon after the 20th week. By the 30th week, AFV is about half the fetal volume, and at term, it is about a quarter of it.¹ The rate of change in AFV is a strong function of gestational age, but in normal pregnancies, a wide volume range is seen, particularly during the second half of gestation; while the AFV range for a given gestational age is very broad between pregnancies, what is most important in clinical practice is not to know exactly its mean variations and their respective reference ranges but to apply a reproducible method enabling detection of those AFV variations.^{2–4}

Detection of abnormalities in AFV is an important part of fetal well-being assessment, because this can correlate with adverse perinatal outcomes: Oligohydramnios is associated

with fetal growth restriction due to placental insufficiency and can be due to rupture of membranes or structural anomalies, particularly of the urinary tract.^{2,5} Polyhydramnios is associated with maternal conditions, such as diabetes and fetal abnormalities, including neural tube defects and gastrointestinal tract obstruction.⁶

Assessment of AFV using ultrasound (US) is widely accepted as the method of choice, as it is easily accessible and safe. Operator experience remains important, and the estimation may also be affected by fetal position and possible transient changes because of the normal dynamic variation of AFV. This means that US assessment is less accurate than invasive methods such as dye dilution techniques.⁷

The two common semiquantitative methods to measure AFV with US are the amniotic fluid index (AFI) and single deepest vertical pool (SDVP), also known as the maximal vertical pocket.

When compared with dye dilution techniques, there is no consensus as to whether AFI is more accurate than SDVP and vice versa.^{2,8} There is also no consensus as to whether AFI or SDVP is the more reproducible method for measuring AFV^{2,5,7,9,10} throughout gestation. Therefore, choice is currently on the basis of clinical preference or local protocols.

The aim of this study is to establish and compare the reproducibility of the two most commonly used methods, AFI and SDVP, throughout gestation.

METHODS

The International Fetal and Newborn Growth Standards for the 21st Century (INTERGROWTH-21st) is an international, multicenter, observational project of fetal and newborn growth currently underway in eight hospitals across the world. All recruited women have low risk pregnancies that fulfill well-defined and strict inclusion criteria at recruitment, details of which have been published elsewhere.¹¹ In the Fetal Growth Longitudinal Study, serial fetal growth scans are performed every 5 ± 1 weeks from 14⁺⁰ to 42⁺⁰ weeks. All US scans are performed using the same commercially available US machine (Philips HD-9, Philips Ultrasound, and Bothell, WA, USA) with curvilinear abdominal transducers (C5-2, C6-3, V7-3). For the purposes of the study, the machine software was engineered to ensure that the measurement values do not appear on screen during the examination.¹² All sonographers taking part in the study underwent standardization¹³ and were subject to quality control processes.¹⁴ The INTERGROWTH-21st protocol was approved by the Oxfordshire Research Ethics Committee. All pregnant participants gave written informed consent.

For this reproducibility study, consecutive pregnant women in the Oxford arm of the INTERGROWTH-21st project were invited to take part in a reproducibility study for assessment of the AFI and SDVP. In order to assess measurement

reproducibility with an equal distribution throughout gestation, we included at least five cases per gestational week from 14 weeks onwards. Three experienced sonographers performed the US scans in observer pairs. Every examination was allocated to an observer pair randomly using a computer-generated randomization algorithm. Each sonographer performed the AFI and SDVP measurement twice, blindly, and independently of the second sonographer. For the intraobserver measurements, each sonographer performed subsequent measurements after completing a full set at the same sitting but at a different time of the exam.

During measurement, the US transducer was held vertical to the uterine contour onto the abdomen and parallel to the maternal sagittal plane. Each amniotic fluid measurement was free of fetal extremities and the umbilical cord² as this may overestimate the volume.¹⁵

The use of color Doppler to ensure absence of the umbilical cord was allowed but not dictated, and was at the discretion of the operator.

The AFI is the sum total of the deepest vertical pockets of liquor in each of the four quadrants into which the uterus is divided by using the linea nigra and above the umbilicus in 24 weeks as proposed by Gramellini *et al.*⁷ If the uterine fundus was below the umbilicus, the uterus was divided into upper and lower halves by a point midway between the symphysis pubis and the top of the uterine fundus. The measurements for the AFI and the SDVP were carried out at the same time. Fetal presentation was recorded as cephalic or noncephalic; fetal activity was also rated as active, quiet, or unable to comment.

As described, the US machine was modified so that measurement values do not appear on screen in order to avoid 'expected value' bias. In addition, the display of caliper placement was removed from the screen before the subsequent sonographer. Data were then extracted from the hard drive of each respective US machine.

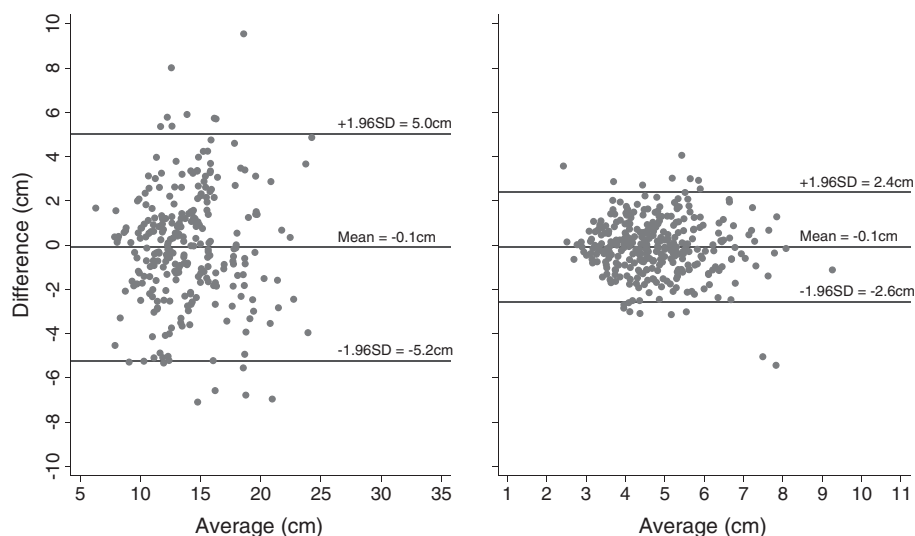


Figure 1 Intraobserver variability for amniotic fluid index (AFI) and single deepest vertical pool (SDVP). Intraobserver variability of AFI (left panel) and SDVP (right panel) expressed as absolute values

STATISTICAL ANALYSIS

Intraobserver and interobserver differences in centimeters were calculated. Measurement differences were also expressed in percentage terms, where the actual difference in centimeter was divided by the mean measurement in centimeter (cm) and multiplied by 100. Both the actual and percentage differences were plotted against the mean measurement, using the method described by Bland and Altman,¹⁶ and 95% limits of agreement were calculated. In addition, within-subject coefficients of variation for SDVP and AFI were calculated.¹⁷ In order to ascertain the effect of fetal presentation and fetal activity on observer agreement, measurement differences were

compared between cephalic and noncephalic fetuses and also between active and quiet fetuses. All analyses were performed using STATA version 11 (Statacorp, College Station, Texas, USA).

RESULTS

We included 175 scans equally distributed between 15 and 41 weeks of gestation. This included 1400 measurements. The participants' mean age was 29.4 years [standard deviation (SD) 4.0], and their mean BMI was 23.2 (SD 2.9).

Mean measurement differences, both within the same observer and between observers, were almost zero, confirming the lack of systematic bias (Figures 1–4). For intraobserver reproducibility

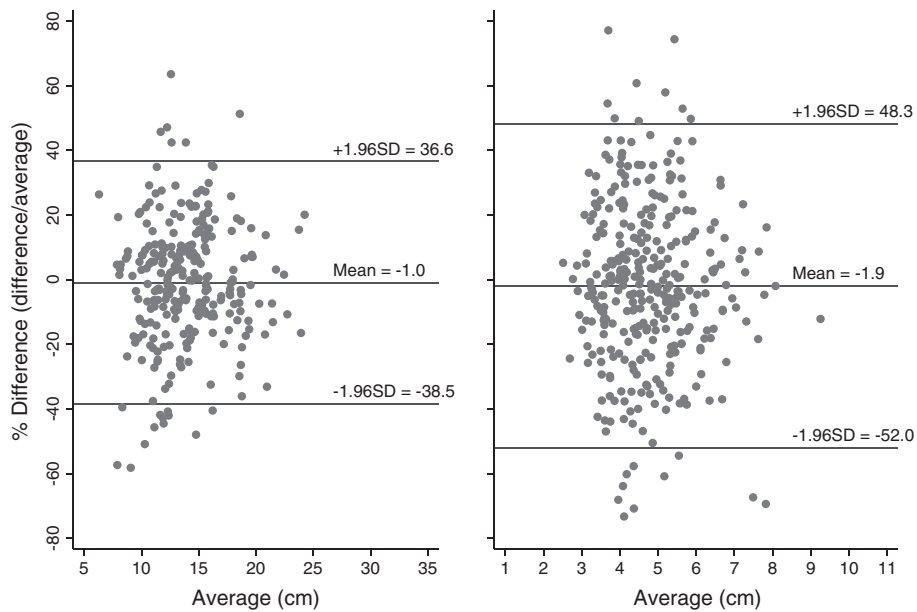


Figure 2 Percentage intraobserver variability for amniotic fluid index (AFI) and single deepest vertical pool (SDVP). Intraobserver variability of AFI (left panel) and SDVP (right panel) expressed as percentage values

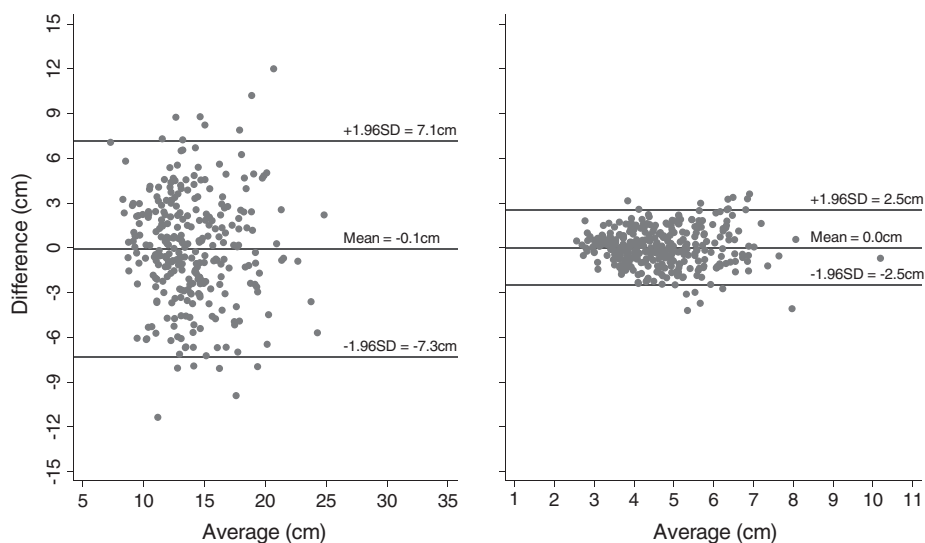


Figure 3 Interobserver variability for amniotic fluid index (AFI) and single deepest vertical pool (SDVP). Interobserver variability of AFI (left panel) and SDVP (right panel) expressed as absolute values

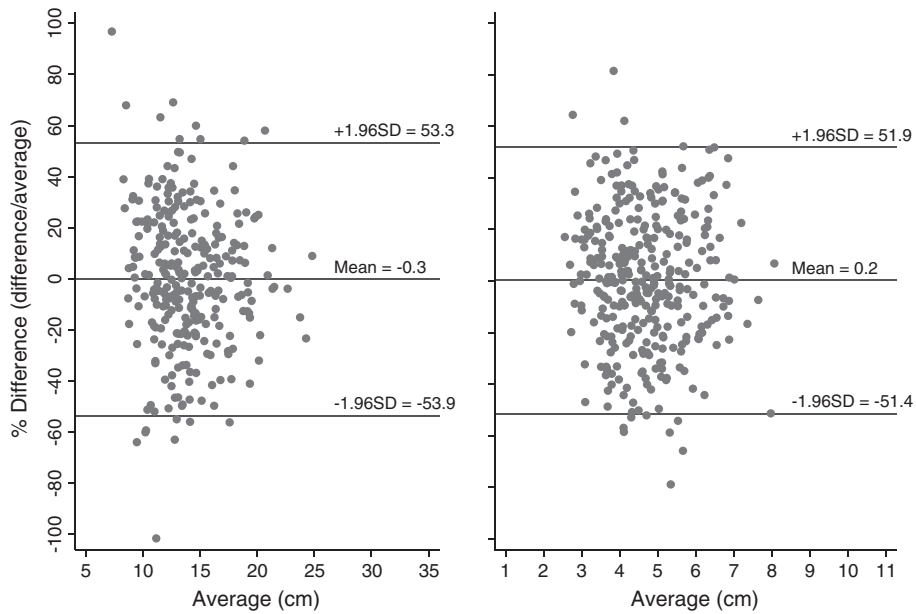


Figure 4 Percentage interobserver variability for amniotic fluid index (AFI) and single deepest vertical pool (SDVP). Interobserver variability of AFI (left panel) and SDVP (right panel) expressed as percentage values

Table 1 Summary of intraobserver and interobserver variability of amniotic fluid index (AFI) and single deepest vertical pool (SDVP)

	Absolute variability (cm)	Percentage variability (%)
Intraobserver variability of AFI and SDVP (+/-1.96 SD)		
AFI	±5	±36
SDVP	±2	±50
Interobserver variability of AFI and SDVP (+/-1.96 SD)		
AFI	±7	±53
SDVP	±3	±52

of AFI, the limits of agreement were -5.2 to 5 cm or -39% to +37%; whereas for SDVP, these were -2.6 to 2.4 cm or -52% to 48%. The interobserver reproducibility of AFI was -7.3 to 7.1 cm or -54% to 53%, and for SDVP, this was -2.5 cm to 2.5 cm or -51% to 52% (Table 1).

Intraobserver coefficient of variation for SDVP was 14% and for AFI 19%, whereas the interobserver coefficient was 19% for both AFI and SDVP.

Univariate analysis of the effect of fetal activity revealed no significant difference between active and quiet babies. Similarly, there was no significant difference between cephalic and noncephalic presentation.

DISCUSSION

Main findings

Our study evaluated the intraobserver and interobserver agreement of measurement for AFI and SDVP in normal, uncomplicated pregnancies with normal AFV. In our study, we established that the intraobserver and interobserver variation for AFI were approximately ±5 and ±7 cm, respectively. For SDVP, the intraobserver and interobserver variation were ±2

and ±3 cm. These values are high. It is difficult to directly compare the reproducibility of AFI and SDVP in actual measurement values, because AFI is an index consisting of four measurements and therefore larger than the single measurement that constitutes SDVP; in addition, both measurements change throughout gestation. To try and overcome this, we expressed the values as percentages. Through this transformation, direct comparisons can be inferred more easily. The interobserver limits of agreement were around ±50% for both SDVP and AFI; intraobserver variation of AFI appeared marginally better at ±38%. We also analyzed potential factors that could lead to increased measurement variation such as fetal activity and fetal presentation. We found no significant effect.

The effect of gestational age on the reproducibility of measurement of AFI and SDVP is not well established. Our study benefits from a large sample size and equal distribution of scans from 14 to 42 weeks of gestation; visual inspection of the Bland-Altman plots confirms that there is not a substantial effect of advancing gestational age on measurement error.

Strengths and limitations

There are several strengths in our study. We examined a large number of subjects and ensured an equal distribution throughout gestation. Formal blinding of all observers was ensured by using an US machine that was modified to ensure no measurement values are visible at the time of caliper placement. A further advantage of this study is the different time points of measurements that allowed assessment of both intra and interobserver differences of the measurements.

This study has some limitations. The pregnancies were uncomplicated, with normal AFV; women were not selected on the basis of normal amniotic fluid, rather on the optimal potential for normal fetal growth. Although we cannot be certain that reproducibility for fetuses with oligohydramnios or

polyhydramnios would be different, we believe that the same factors that affect reproducibility of normal amniotic fluid affect assessment of abnormal amniotic fluid. It is not common practice to assess reproducibility in different subgroups (e.g. reproducibility of fetal growth in normal growth vs growth restricted fetuses). Nevertheless, it is possible that the measurement error, when AFV is abnormal, is not the same as that to normal and that measurement variation increases with liquor volume.

Unlike the study by Gramellini *et al.*,² our data are based on single measurements, which reflect common clinical practice and may therefore be more relevant. Other notable conditions were that observers were well trained in US and had ample time to complete the examinations. It is possible that under different conditions reproducibility may be different; however, our aim was to report reproducibility under optimal conditions to serve as a standard.

Interpretation

Despite the universal use of US to estimate AFV and the implications of an abnormal measurement, there are still questions as to which method is the most reliable.^{8,15} Previous studies have assessed the variation of AFI and SDVP^{2,8,15,18} but most assessed the intraobserver and interobserver variation of either AFI or SDVP measurement alone or used small sample sizes or limited range of gestational ages. Other researchers have focused on how well AFI and SDVP correlate with actual AFV, using dye dilution techniques as the gold standard; these studies were carried out on near term pregnancies or patients who subsequently underwent amniocentesis^{4,7,10} and in general suggest that the correlation is poor to moderate. Other studies have assessed which one of the two methods is most accurate for predicting perinatal morbidity and mortality.^{2,9,19–22} A meta-analysis²³ comparing AFI and SDVP in preventing adverse perinatal outcome concluded that the SDVP may be better than AFI, because it is associated with a lower rate of labor induction. Similarly, it has been shown that the use of SDVP may be more accurate in the prediction of oligohydramnios in late pregnancy,^{18,24–26} as AFI in the third trimester may overestimate the incidence of oligohydramnios.^{7,9}

Our finding that intraobserver and interobserver reproducibility for SDVP and AFI is very similar combined with clinical data suggesting that AFI may lead to increased intervention without benefit suggests that SDVP may be a preferable method until a better tool for assessing AFV is developed.^{27–29}

CONCLUSION

In this study, we presented the intraobserver and interobserver variation of AFI and SDVP measurement, both in actual measurement difference and percentage difference values. The limits of agreement are wide for both AFI and SDVP, and none is consistently superior to the other.

The choice of measurement for amniotic fluid measurement should take into account both the measurement reproducibility, and its ability to predict perinatal outcome. Our study has demonstrated that intraobserver and interobserver reproducibility

for both measurements is quite high throughout gestation. Further work is warranted in order to standardize a reproducible method to assess AFV. Until then, care should be taken in clinical practice in the interpretation of an abnormal AFV value.

ACKNOWLEDGEMENTS

We are grateful to the pregnant volunteers who patiently underwent the US scans for this exercise and to Fenella Roseman and Annie Laister for recruiting them.

Joyce Sande is affiliated to both the Aga Khan University Hospital Nairobi and the INTERGROWTH-21st study. Ippokratis Sarris is affiliated to Newcastle Fertility Center at Life UK.

We would like to thank Philips Healthcare for providing the specifically modified HD9 US machines and technical assistance.

Ethics approval

The INTERGROWTH-21st protocol was approved by the Oxfordshire Research Ethics Committee C, reference: 08/H0606/139.

Authorship

J. S. analyzed and interpreted the data and wrote the manuscript. C. I., I. S., and A. P. conceived and designed the study and also contributed to interpretation of the data. E. O. conducted all the analyses and contributed to the interpretation of the data. All authors revised the article critically and edited and approved the final version.

WHAT'S ALREADY KNOWN ABOUT THIS TOPIC?

- Amniotic fluid is a key indicator of fetal well-being in the second half of gestation. Detection of abnormal amniotic fluid volume (AFV) is significant because it often correlates with adverse outcomes. Measurement of AFV can be carried out via invasive or noninvasive techniques. Ultrasound is widely accepted as the method of choice clinically, as it is easily accessible and safe. However, it is operator dependent, semiquantitative, and less accurate than invasive methods such as dye dilution techniques. There are two common ways to measure AFV with ultrasound: the amniotic fluid index (AFI) and single deepest vertical pool (SDVP). There is no consensus as to whether AFI or SDVP is the more reproducible method for measuring AFV throughout gestation. Therefore, choice is currently on the basis of clinical preference or local protocols.
- The aim of this study is to establish and compare the reproducibility of the most commonly used two methods, AFI and SDVP, throughout gestation.

WHAT DOES THIS STUDY ADD?

- This is a large study of 175 fetuses and 1400 measurements. In this study, we presented the intraobserver and interobserver variation of AFI and SDVP measurement, both in actual measurement difference and percentage difference values. The limits of agreement are wide for both AFI and SDVP and none is consistently superior to the other. Further work is warranted in order to standardize a reproducible method to assess AFV. Until then, care should be taken in clinical practice in the interpretation of an abnormal AFV value.

REFERENCES

1. Modena AB, Fieni S. Amniotic fluid dynamics. *Acta Biomed* 2004;75 (Suppl 1):11–3.
2. Gramellini D, Fieni S, Verrotti C, *et al.* Ultrasound evaluation of amniotic fluid volume: methods and clinical accuracy. *Acta Biomed* 2004;75(Suppl 1):40–4.
3. Croom CS, Baniyas BB, Ramos-Santos E, *et al.* Do semiquantitative amniotic fluid indexes reflect actual volume? *Am J Obstet Gynecol* 1992;167:995–9.
4. Magann EF, Nolan TE, Hess LW, *et al.* Measurement of amniotic fluid volume: accuracy of ultrasonography techniques. *Am J Obstet Gynecol* 1992;167:1533–7.
5. Sherer DM, Langer O. Oligohydramnios: use and misuse in clinical management. *Ultrasound Obstet Gynecol* 2001;18:411–9.
6. Dashe JS, McIntire DD, Ramus RM, *et al.* Hydramnios: anomaly prevalence and sonographic detection. *Obstet Gynecol* 2002;100:134–9.
7. Nabhan AF, Abdelmoula YA. Amniotic fluid index versus single deepest vertical pocket as a screening test for preventing adverse pregnancy outcome (Cochrane Review). JohnWiley & Sons, Ltd: Oxford, U.K., 2009.
8. Bruner JP, Reed GW, Sarno AP, *et al.* Intraobserver and interobserver variability of the amniotic fluid index. *Am J Obstet Gynecol* 1993;168:1309–13.
9. Verrotti C, Bedocchi L, Piantelli G, *et al.* Amniotic fluid index versus largest vertical pocket in the prediction of perinatal outcome in post-term pregnancies. *Acta Biomed* 2004;75(Suppl 1):67–70.
10. Magann EF, Chauhan SP, Barrilleaux PS, *et al.* Amniotic fluid index and single deepest pocket: weak indicators of abnormal amniotic volumes. *Obstet Gynecol* 2000;96:737–40.
11. Villar J, Altman D, Purwar M, *et al.* The objectives, design and implementation of the INTERGROWTH-21st Project. *BJOG* 2013;120:9–26.
12. Papageorgiou AT, Sarris I, Ioannou C, *et al.* Ultrasound methodology used to construct the fetal growth standards in the INTERGROWTH-21st Project. *BJOG* 2013;120(Suppl. 2): 27–32.
13. Sarris I, Ioannou C, Dighe M, *et al.* Standardization of fetal ultrasound biometry measurements: improving the quality and consistency of measurements. *Ultrasound Obstet Gynecol* 2011;38:681–7.
14. Sarris I, Ioannou C, Ohuma E, *et al.* Standardisation and quality control of ultrasound measurements taken in the INTERGROWTH-21st project. *BJOG* 2013;120(Suppl. 2):33–7.
15. Peedicayil A, Mathai M, Regi A, *et al.* Inter- and intra-observer variation in the amniotic fluid index. *Obstet Gynecol* 1994;84:848–51.
16. Altman DG. Measurement in medicine: the analysis of method comparison studies. *Statistician* 1983;32:307–17.
17. Shoukri MM, Colak D. Comparison of two dependent within subject coefficients of variation to evaluate the reproducibility of measurement devices. *BMC Med Res Methodol* 2008;8 (24):1471–2288
18. Halperin ME, Fong KW, Zalev AH, *et al.* Reliability of amniotic fluid volume estimation from ultrasonograms: intraobserver and interobserver variation before and after the establishment of criteria. *Am J Obstet Gynecol* 1985;153:264–7.
19. Moses J, Doherty DA, Magann EF, *et al.* A randomized clinical trial of the intrapartum assessment of amniotic fluid volume: amniotic fluid index versus the single deepest pocket technique. *Am J Obstet Gynecol* 2004;190:1564–70.
20. Rutherford SE, Phelan JP, Smith CV, *et al.* The four-quadrant assessment of amniotic fluid volume: an adjunct to antepartum fetal heart rate testing. *Obstet Gynecol* 1987;70:353–6.
21. Manning FA, Platt LD, Sipos L. Antepartum fetal evaluation: development of a fetal biophysical profile. *Am J Obstet Gynecol* 1980;136:787–95.
22. Chauhan SP, Magann EF, Perry KGJ, *et al.* Intrapartum amniotic fluid index and two-diameter pocket are poor predictors of adverse neonatal outcome. *J Perinatol* 1997;17:221–4.
23. Nabhan AF, Abdelmoula YA. Amniotic fluid index versus single deepest vertical pocket: a meta-analysis of randomized controlled trials. *Int J Gynaecol Obstet* 2009;104:184–8.
24. Seffah JD, Armah JO. Amniotic fluid index for screening late pregnancies. *East Afr Med J* 1999;76:348–51.
25. Rutherford SE, Smith CV, Phelan JP, *et al.* Four-quadrant assessment of amniotic fluid volume. Interobserver and intraobserver variation. *J Reprod Med* 1987;32:587–9.
26. Magann EF, Kinsella MJ, Chauhan SP, *et al.* Does an amniotic fluid index of ≤ 5 cm necessitate delivery in high-risk pregnancies? A case-control study. *Am J Obstet Gynecol* 1999;180:1354–9.
27. Magann EF, Doherty DA, Chauhan SP, *et al.* How well do the amniotic fluid index and single deepest pocket indices (below the 3rd and 5th and above the 95th and 97th percentiles) predict oligohydramnios and hydramnios? *Am J Obstet Gynecol* 2004;190:164–9.
28. Alfirevic Z, Luckas M, Walkinshaw SA, *et al.* A randomised comparison between amniotic fluid index and maximum pool depth in the monitoring of post-term pregnancy. *Br J Obstet Gynaecol* 1997;104:207–11.
29. Chauhan SP, Doherty DD, Magann EF, *et al.* Amniotic fluid index vs single deepest pocket technique during modified biophysical profile: a randomized clinical trial. *Am J Obstet Gynecol* 2004;191:661–6.