Published in final edited form as:

Am J Perinatol. 2012 October ; 29(9): . doi:10.1055/s-0032-1314887.

Comparative Analysis of 2-D Versus 3-D Ultrasound Estimation of the Fetal Adrenal Gland Volume and Prediction of Preterm Birth

Ozhan M. Turan, M.D., PhD¹, Sifa Turan, M.D., RDMS¹, Irina A. Buhimschi, M.D.², Edmund F. Funai, M.D.², Katherine H. Campbell, M.D.², Ozan M. Bahtiyar, M.D.², Chris R. Harman, M.D.¹, Joshua A. Copel, M.D.², Ahmet A Baschat, M.D.¹, and Catalin S. Buhimschi, M.D.²

¹University of Maryland School of Medicine, Dept. Obstet. Gynecol. and Reprod. Sciences

²Yale University School of Medicine, Dept. Obstet. Gynecol. and Reprod. Sciences

Abstract

Objective—We aim to test the hypothesis that 2D fetal AGV measurements offer similar volume estimates as volume calculations based on 3D technique

Methods—Fetal AGV was estimated by 3D ultrasound (VOCAL) in 93 women with signs/symptoms of preterm labor and 73 controls. Fetal AGV was calculated using an ellipsoid formula derived from 2D measurements of the same blocks ($0.523 \times \text{length} \times \text{width} \times \text{depth}$). Comparisons were performed by intra-class correlation coefficient (ICC), coefficient of repeatability, and Bland-Altman method. The cAGV (AGV/fetal weight) was calculated for both methods and compared for prediction of PTB within 7 days.

Results—Among 168 volumes, there was a significant correlation between 3D and 2D methods ($\text{ICC}=0.979[95\% \text{CI}: 0.971-0.984]$). The coefficient of repeatability for the 3D was superior to the 2D method (Intra-observer 3D: 30.8, 2D:57.6; inter-observer 3D: 12.2, 2D: 15.6). Based on 2D calculations, a cAGV $433\text{mm}^3/\text{kg}$, was best for prediction of PTB (sensitivity: 75%(95%CI=59-87); specificity: 89%(95%CI=82-94). Sensitivity and specificity for the 3D cAGV (cut-off $420\text{mm}^3/\text{kg}$) was 85%(95%CI=70-94) and 95%(95%CI=90-98), respectively. In receiver-operating-curve curve analysis, 3D cAGV was superior to 2D cAGV for prediction of PTB ($z=1.99$, $p=0.047$).

Conclusion—2D volume estimation of fetal adrenal gland using ellipsoid formula cannot replace 3D AGV calculations for prediction of PTB.

Keywords

3D ultrasound; volume measurement; fetal adrenal gland; preterm birth

Introduction

Preterm labor is one of the most common causes of maternal antenatal hospitalization and precedes 40-50% of preterm births (PTB).^{1,2} Despite significant advances in perinatal and neonatal care, PTB continues to be one of the leading causes of neonatal mortality and morbidity.^{3,4} Clinicians continue to have difficulty in accurately identifying and predicting

the risk for PTB because of the limitations of the available predictive tests used alone or in combination.^{5,6}

The ability of 3-dimensional (3D) fetal adrenal gland volume (AGV) measurements to predict PTB was previously reported by our group.⁷ However, wide implementation of the 3D fetal AGV measurements is hindered by the complexity of this methodology. So, it would be better to have a simpler technique.

The fetal adrenal gland has a discoid shape.⁸ 2D ultrasound can be used to calculate the volume of ellipsoid organs with 10-20% error.^{9,10} Goldstein et al. demonstrated that a prolate ellipsoid formula (volume = 0.52 (length × anteroposterior diameter × transverse diameter) was a suitable volume calculation method for spherical objects.¹¹ Yet, no prior studies have evaluated calculation of AGV based on prolate ellipsoid formula.

The aim of this study was to test the hypothesis that 2D fetal AGV measurements offer similar volume estimates as volume calculations based on 3D technique.

Materials and Methods

Study design

This study is a secondary analysis of data collected as part of a prospective observational study performed at University of Maryland and Yale University Schools of Medicine.¹² The study was conducted from 2005 to 2009 under research protocols approved by Human Investigational Committees at both institutions. All patients provided signed informed consent.

First, 3D ultrasound blocks of the fetal adrenal gland were acquired on 93 consecutive singleton fetuses of women who presented between 18 to 34 weeks of gestation with symptoms of preterm labor and/or preterm premature rupture of membranes (PPROM). Adrenal gland volume acquisition and fetal biometry were performed within 24 hours after admission. Gestational age (GA) was established based on an ultrasonographic examination prior to 20 weeks. Inclusion criteria included symptoms of preterm labor or PPRM **in singleton pregnancies**. Preterm labor was defined as presence of regular uterine contractions with cervical effacement and/or advanced cervical dilatation >3cm. A diagnosis of PPRM was confirmed by visualization of amniotic fluid pooling during a sterile speculum examination, positive nitrazine and ferning tests. Suspected fetal growth restriction (IUGR) (ultrasonographic estimated fetal weight <10 percentile for GA), maternal medical complications (i.e. hypertension, preeclampsia, diabetes, thyroid or adrenal diseases), fetal structural abnormalities diagnosed at the time of assessment or at delivery, presence of fetal heart rate abnormalities at enrollment (bradycardia or prolonged variable decelerations), and **any other complications which required urgent delivery were considered as exclusion criteria**.

Following enrollment, study group women were followed prospectively up to the point of delivery. The delivery interval was defined as the number of days between the enrolment sonogram and delivery of the fetus. The clinical management of the patients was left to the discretion of the clinical team and carried out according to the American College of Obstetrics and Gynecology recommendations.¹

To provide GA reference values for the AGVs, we evaluated 73 consecutive singleton fetuses carried by asymptomatic healthy pregnant women who presented to the low risk clinic for a routine prenatal care visit. Similar to the study group, all the control patients were followed longitudinally up to the point of delivery.

Ultrasound measurements and calculation of the AGVs

Fetal adrenal gland volume acquisition was performed with the use of the Voluson 730 and E8 systems (Voluson Expert, General Electric Medical Systems, Milwaukee, WI, USA), equipped with a 4-8 MHZ curved array transducer as previously described.⁷ Three dimensional block was obtained from **only one adrenal gland**, as a rule the gland closest to the probe was chosen. Three-dimensional blocks were analyzed offline by a single investigator (OMT) that was blinded to the pregnancy outcome. Calculation of AGV was performed using VOCAL (Virtual Organ Computer-aided AnaLysis, 4D View; General Electric Medical System, Milwaukee, WI, USA) software package as previously described.⁷ Three different volumes were calculated from each patient and recorded. The mean of 3 measurements **was** used for calculation of corrected adrenal gland volume (cAGV). The cAGV, (variable independent of GA) was calculated by dividing the acquired AGV to the ultrasound estimated fetal weight (EFW).

Next, the dimensions of the fetal adrenal gland were measured within the same blocks used for prior volume calculations. **Some operator who was blinded to previous 3D volume result was measured the 2D dimensions.** Length (L) was measured in transverse or sagittal planes. Width was measured in transverse or coronal planes. Depth was measured in sagittal and coronal planes. All dimensions were measured 3 times in each patient and recorded. The 2D image derived AGVs were computed 3 times from each patient by using the prolate ellipsoid formula ($0.523 \times \text{length} \times \text{width} \times \text{depth}$), as previously described.¹¹ **The mean of 3 measurements was used for final analysis.**

Data analysis and statistical methodology

Kolmogorov-Smirnov test was used to verify distribution of each data. Categorical variables were analyzed using ² or Fisher's exact test according their size. Continuous variables were analyzed using Mann-Whitney or Student t-test according their distributions. Wilcoxon Signed Ranks test was used for comparison of paired samples. Regression analysis was used for covariate analysis. Pearson correlation test was used to investigate correlation between 2D and 3D volume calculation methods. A coefficient of correlation (r) result close to 1 was accepted as a good correlation.

The agreement between 2D versus 3D methodology for AGV calculation was estimated by comparing the mean of three consecutive measurements of the AGV obtained by using each of the two methods. For clinical relevance, knowledge of the agreement between replicate measurements (precision) is critical. The agreement among fetal AGV measurements evaluated using the 2D and 3D methods were estimated by calculating the intra-class correlation coefficient (ICC), as previously recommended.¹³ Intra-class correlation coefficient score close to 1 was accepted as an ideal result. We performed a Bland-Altman analysis of the differential value between the two methodologies to assess if the differences between 3D and 2D techniques were significant and whether the two methods may be used interchangeably.¹⁴ For this analysis, percentage difference in mean and standard error of mean (SEM) were calculated. Data were presented as mean percent difference and 95% confidence limits of agreement (mean $\pm$ 95% confidence limits of agreement). Confidence limits of agreement represent lower and upper limits of 2SD of mean differences as percentage. Standard error of mean close to 0 was accepted as an ideal result.¹⁰ Coefficient of repeatability was calculated as $\text{SEM} \times 2$. This value represents $2 \times$ the standard deviation of change of mean in Bland-Altman plot. Ideal precision was defined as a small change in the mean (approaching zero) and narrow coefficient of repeatability.

Fetal AGVs calculated based on both 2D and 3D methods were analyzed in relationship to the time interval from the moment of AGV data acquisition to delivery. The ability of 2D

and 3D AGVs to predict PTB within 7 days was investigated using receiver operating curve (ROC) analysis to be able to identify their clinical utility. Sensitivity, specificity, positive (+LR) and negative (-LR) likelihood ratios were calculated. A LR of 1.0 indicates no change from pre-test probability, suggesting a useless test without any predictive value. As general guidelines, a +LR>2, corresponds to the probability of disease of approximately 67%. Values of +LR >5 argue strongly for presence of the sought condition, while values of LR(-) <0.2 strongly **militated** against its presence^{15,16}. A higher +LR and lower -LR was accepted as a good test performance.¹⁷ In order to test intra-observer variability, the first and third measurements of 2D and 3D ultrasound derived volumes were compared in all population. For inter-observer variability, 20 patients were randomly selected. Two different investigators (MOT and ST) calculated the AGV using 2D and 3D methods. Intra-class correlation coefficient and Bland-Altman method were used for analysis of intra- and inter-observer variability in 2D and 3D calculation methods.

SPSS 11 (SPSS Inc., Chicago, IL, USA), MedCalc (Broekstraat, Belgium) and Microsoft Excel were used for data analysis. This study was powered to calculate 3D and 2D volume differences ($\alpha=0.05$, $\beta=0.1$, Null hypothesis=4%, difference=10%, $n=165$). Ninety percent of the AGV blocks were used in our previous publications and were re-analyzed for this study.^{7,12} Data acquired from sixteen new patients were also included.

Results

Total of 182 singleton pregnant women were recruited for the study. Adrenal gland blocks were successfully measured in 166 (91%) of them. Acquisition of 3D blocks was performed before first dose of betamethasone injection in 78 (84%) of the cases with signs of preterm labor (PTL)/PPROM. None of the blocks were acquired after 2nd dose of betamethasone injection. Demographic and outcome variables of study population are presented in Table 1. As expected, previous history of PTB and presence of PPRM were more common in the PTL/PPROM group. Patients in this group were delivered earlier and had lower birth-weights, on average. The majority of the patients in the PTL/PPROM group delivered before 37 weeks and almost half of them delivered within 7 days of presentation. Gestational age at enrollment was significantly correlated with both 2D and 3D derived volume measurements (Pearson's $r=0.72$ and 0.73 , $p<0.0001$ respectively). Six fetuses of the control group were delivered preterm. At the time of enrolment their AGVs were not estimated as enlarged by either method.

Inter-method agreement

ICC analysis showed there was a high degree of reliability among 3D and 2D volume calculations (ICC= 0.979 [95%CI: 0.971-0.984]). However, 2D calculation of the AGVs showed larger volumes than 3D derived measurements (mean $\pm$ SEM difference: $1.3\% \pm 20.9\%$) (Wilcoxon Signed Ranks test $p<0.03$), with a coefficient of repeatability that reached 41.8% (Table 2, Figure 1). This result implies that 2D estimation of the AGV may be up to $\pm 40\%$ different than 3D estimation. Such difference was particularly prominent for smaller AGVs. (Figure 2a and 2b). Therefore, we appreciated that estimation of the fetal AGV using the 2D technique has a large margin of error.

Prediction of preterm birth by using 2D versus 3D estimation of the AGVs

The sensitivity, specificity, +LR and -LR of cAGV based on 3D and 2D methods are listed in Table 3. The 3D ultrasound derived AGV volume was better than the 2D method in predicting PTB within 7 days ($z=1.99$, $p=0.047$) (Figure 3).

Intra-observer agreement

A high degree of reliability was found between the 1st and 3rd calculations of 3D (ICC=0.982 (95% CI: 0.976-0.987) and 2D (ICC=0.925 (95% CI: 0.899-0.945) measurement of the fetal AGVs. Figure 4a displays the Bland-Altman plot for the percent change of the mean differences and 95% limits of agreement for the intra-observer measurements of 3D (mean, 0.7; 95% limits of agreement, -29.4 to 30.9%) and 2D (Figure 4b) (mean, -2.65; 95% limits of agreement, -49.3 to 44%) ultrasound derived volume calculations. Coefficient of repeatability was smaller in 3D than 2D volume calculations (30.8% vs 57.6%) (Table 2). These results show that 3D calculation of fetal adrenal gland volume has higher intra-observer reliability and reproducibility than 2D calculations.

Inter-observer agreement

A high degree of inter-observer reliability was found between 3D (ICC=0.998 (CI: 0.994-0.999)) and 2D (ICC=0.997 (CI: 0.992-0.999)) methods (Table 3). Figure 5a displays the Bland-Altman plot for the percent change of the mean differences and 95% limits of agreement for the inter-observer measurements of the AGVs by using the 3D (mean, -0.5; 95% limits of agreement, -12.5 to 11.6%) and 2D (mean, -6.9; 95% limits of agreement, -22.2 to 8.4%) (Figure 5b) volume calculation. Coefficient of repeatability was smaller in 3D than 2D volume calculations (24.4% vs 31.2%) (Table 2). These results show 3D volume calculation for fetal adrenal gland has higher inter-observer reliability and reproducibility than 2D calculations.

Discussion

Evaluation of the fetal AGV can improve the clinicians' ability to predict PTB. While 3D ultrasound technology is increasingly utilized in clinical perinatology, estimation of the fetal AGV continues to require specialized expertise. Studies indicate that 3D ultrasound does not introduce significant errors compared to 2D ultrasound and that 3D ultrasound offers improved accuracy compared with the traditional 2D method.¹⁸ Estimation of kidney volume using 3D and 2D ultrasound have shown that 3D ultrasound allows for accurate volume measurements.^{10,19} Chou et al. reported that estimation of cervical carcinomas' volume using 3D technology was more precise than evaluation of the tumors' volume using 2D technique.²⁰ Similarly, 3D ovarian volume estimations were found to be more accurate than 2D volume calculations.²¹

This is the first study to compare the precision of individual estimates of fetal AGV using both 2D and 3D methods. We found excellent ICC for comparison of 3D and 2D volume measurement. The 2D methodology estimated larger AGVs with a relatively large SEM as compared to 3D volume estimates. This implies decreased precision of the 2D volume measurement. Precision of a test depends on the study population variance. If the precision value lies within a broad range, the calculated SEM will be relatively large, and thus imprecise.

Riccabona et al showed that the accuracy of a volume measurement applicable to ellipsoid shaped objects obtained similar whether the volume is estimated by either 3D or 2D ultrasound. However, 3D ultrasound measures irregularly shaped object more accurately.^{22,23} In this study we found that 2D volume derived cAGV had a significantly lower predictive value than 3D volume based cAGV. There is a disproportional increase in the dimensions of the adrenal gland when this organ becomes enlarged. In our most recent study we found that depth of the adrenal gland correlates best with PTB.¹² This observation may be related to the shape of the fetal adrenal gland which becomes more irregular with

increasing size and makes determination of the AGVs based on 2D calculations less accurate.

In our study, ICC for inter- and intra-observer agreements for 3D and 2D volume demonstrated excellent agreement, but 2D volume calculations showed higher SEM. Bland-Altman coefficient of repeatability is closely related to the SEM of a technique, and allows for a precise estimation of how close two measurements taken from the same scan are likely to be.¹⁴ Similar to our study, Brett et al. reported high degree of ICC of intra- and inter-observer variability for ovarian volume measurements.²⁴ However, the authors also proved that both methods were lacking from good inter- and intra- observer variability, particularly when estimating low ovarian volumes. In our study, 2D intra-observer variability was particularly poor when measuring smaller adrenal glands.

Three-dimensional versus 2D estimation of the kidney volume showed that length, width and depth measurement of the kidney had higher intra- and inter- observer variability.^{10,25} The authors reasoned that this is due to difficulty in obtaining optimal and reproducible image planes during each attempt. Difficulty in obtaining the exact same views of the fetal adrenal gland may also explain the large SEM observed by using the 2D methodology. This problem can be alleviated by measuring adrenal gland dimension and fetal zone in the same 2D view, and using their ratios as a marker of adrenal gland enlargement.¹² Although repeat exams might not be at the same level, the ratio should remain similar.

Our results suggest that 3D volume measurement is better than 2D volume calculation for prediction of impending PTB. Still, sensitivity and specificity of 2D volume calculation are comparable to other PTB assessment methods which are currently used in our clinical practice.^{26,27} Yet, individual measurements necessary to derive the fetal AGV by using 2D calculation are not optimally precise, and thus may have limited clinical value. Bernhard et al. compared the organ volume calculation with the use of a similar ellipsoid formula. Similar to us, the authors concluded that 2D based volumes are less correlated with the actual volumes.¹⁰

This study had enough power to identify volume differences when 3D and 2D calculation of the fetal AGV are used. However, we recognize our limitation. The 2D measurements of fetal adrenal gland were obtained in acquired 3D blocks. This may create bias. On the other hand, using same blocks to obtain 3D volume calculation and to measure dimensions has a potential advantage to compare accuracy of both volume calculation methods. Since 3D volume estimation of adrenal gland performs better than 2D volume measurement which was derived from 3D blocks, we may extrapolate that using 2D ultrasound for volume calculations will not be useful. **In our most recent publication we reported that demonstration of enlargement in depth of fetal zone by 2D ultrasound is a better predictor of PTB than 3D volume calculation¹².** This method is easier than volume measurement and readily available. Since the 2D volume measurement performs worse than 3D volume measurement even when we used same blocks, we opine that the best diagnostic modality is estimation of the fetal adrenal gland zone. In 16% of the cases adrenal gland volumes were measured after first dose of the betamethasone. Fetal AGV were measured 24 and 48 hours after first dose of betamethasone in a subset of patients. We did not find any statistical differences in AGVs.

In summary, 2D volume estimation of the fetal adrenal gland cannot replace 3D AGV calculations for prediction of PTB.

Acknowledgments

No relevant commercial interest.

References

1. ACOG practice bulletin. Management of preterm labor. Number 43, May 2003. *Int J Gynaecol Obstet.* 2003; 82:127–35. [PubMed: 12834934]
2. Liu S, Heaman M, Sauve R, Liston R, Reyes F, Bartholomew S, Young D, Kramer MS, Maternal Health Study Group of the Canadian Perinatal Surveillance System. An analysis of antenatal hospitalization in Canada, 1991–2003. *Matern Child Health J.* 2007; 11:181–7. [PubMed: 17089198]
3. Heron M, Tejada-Vera B. Division of Vital Statistics Deaths: Leading cause of 2005. *National Vital Statistic Reports.* 2009; 58:1–100.
4. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Mathews TJ, Kirmeyer S, Osterman MJ. Division of Vital Statistics. Births: Final data for 2007. *National Vital Statistic Reports.* 2010; 58:1–85.
5. Iams J. Prediction and early detection of preterm labor. *Obstet Gynecol.* 2003; 101:402–12. [PubMed: 12576267]
6. Sanchez-Ramos L, Delke I, Zamora J, Kaunitz AM. Fetal fibronectin as a short-term predictor of preterm birth in symptomatic patients: a meta-analysis. *Obstet Gynecol.* 2009; 114:631–40. [PubMed: 19701045]
7. Turan OM, Turan S, Fuani EF, Buhimschi IA, Copel JA, Buhimschi CS. Fetal adrenal gland volume. A novel method for identify women at risk for impending preterm birth. *Obstet Gynecol.* 2007; 109:855–862. [PubMed: 17400846]
8. Droste S, Fitzsimmons J, Pascoe-Mason J, Shepard TH, Mack LA. Size of the fetal adrenal in bilateral renal agenesis. *Obstet Gynecol.* 1990; 76:206–9. [PubMed: 2196498]
9. Ravichandran G, Fellows GJ. The accuracy of a hand-held real time ultrasound scanner for estimating bladder volume. *Br J Urol.* 1983; 55(1):25–7. [PubMed: 6824861]
10. Bernhard PL, Stadler A, Schamp S, Koller A, Voracek M, Heinz G, Helbich TH. Accuracy of volume measurement in human Cadaver kidney. *Investig Radiol.* 2002; 37:489–495. [PubMed: 12218444]
11. Goldstein SR, Horii SC, Snyder JR, Raghavendra BN, Subramanyam B. Estimation of nongravid uterine volume based on a nomogram of gravid uterine volume: its value in gynecologic uterine abnormalities. *Obstet Gynecol.* 1988; 72:86–90. [PubMed: 3288932]
12. Turan OM, Turan S, Funai EF, Buhimschi IA, Campbell CH, Bahtiyar OM, Harman C, Copel JA, Buhimschi CS, Baschat A. Ultrasound measurement of fetal adrenal gland enlargement: An accurate predictor of preterm birth (PTB). *Am J Obstet Gynecol.* 2011; 204:311.e1–311.e10. Epub 2011 Feb. [PubMed: 21292230]
13. Koch, GG. Intraclass correlation coefficient. In: Kotz, Samuel; Johnson, NormanL, editors. *Encyclopedia of Statistical Sciences.* Vol. 4. New York: John Wiley & Sons; 1982. p. 213–217.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1:307–310. [PubMed: 2868172]
15. Foy R, Warner P. About time: diagnostic guidelines that help clinicians. *Qual Saf Health Care.* 2003; 12:205–9. [PubMed: 12792011]
16. Wang ST, Pizzolato S, Demshar HP. Receiver operating characteristic plots to evaluate Guthrie, Wallac, and Isolab phenylalanine kit performance for newborn phenylketonuria screening. *Clin Chem.* 1997; 43:1838–42. [PubMed: 9342001]
17. Attia J. Moving beyond sensitivity and specificity: using likelihood ratios to help interpret diagnostic tests. *Aust Prescr.* 2003; 26:111–13.
18. Gilja OH, Hausken T, Berstad A, Odegaard S. Measurements of organ volume by ultrasonography. *Proc Instn Mech Engrs.* 1999; 213:247–59.
19. Matre K, Stokke EM, Martens D, Gilja OH. In vitro volume estimation of kidneys using three-dimensional ultrasonography and a position sensor. *Eur J Ultrasound.* 1999; 10:65–73. [PubMed: 10502641]
20. Chou CY, Hsu KF, Wang ST, Huang SC, Tzeng CC, Huang KE. Accuracy of three-dimensional ultrasonography in volume estimation of cervical carcinoma. *Gynecol Oncol.* 1997; 66:89–93. [PubMed: 9234927]

21. Kyei-Mensah A, Maconochie N, Zaidi J, Pittrof R, Campbell S, Tan SL. Transvaginal three-dimensional ultrasound: reproducibility of ovarian and endometrial volume measurements. *Fertil Steril*. 1996; 66:718–22. [PubMed: 8893673]
22. Riccabona M, Nelson TR, Pretorius DH. Three-dimensional ultrasound: accuracy of distance and volume measurements. *Ultrasound Obstet Gynecol*. 1996; 7:429–34. [PubMed: 8807760]
23. Wong JB, Gerscovich EO, Cronan MS, Seibert AJ. Accuracy and precision of in vitro volumetric measurement by three dimensional sonography. *Invest Radiol*. 1996; 31:26–9. [PubMed: 8850362]
24. Brett S, Bee N, Wallace WH, Rajkhowa M, Kelsey TW. Individual ovarian volumes obtained from 2-dimensional and 3-dimensional ultrasound lack precision. *Reprod Biomed Online*. 2009; 18(3): 348–51. [PubMed: 19298733]
25. Emamian SA, Nielsen MB, Pedersen JF. Intraobserver and interobserver variations in sonographic measurements of kidney size in adult volunteers. A comparison of linear measurements and volumetric estimates. *Acta Radiol*. 1995; 36:399–401. [PubMed: 7619620]
26. Iams J. Prediction and early detection of preterm labor. *Obstet Gynecol*. 2003; 101:402–12. [PubMed: 12576267]
27. Sanchez-Ramos L, Delke I, Zamora J, Kaunitz AM. Fetal fibronectin as a short-term predictor of preterm birth in symptomatic patients: a meta-analysis. *Obstet Gynecol*. 2009; 114:631–40. [PubMed: 19701045]

Figure 1.

Bland-Altman plot for percentage of mean difference and 95% limits of agreement between 3D and 2D volume calculations.

Figure 2a

Figure 2b

Figure 2.

These graphs show intra-observer relationship between (a) 3D 3rd and 1st volume measurement and (b) 2D 3rd and 1st volume measurement. Data transformed to logarithmic values to achieve normal distribution. Solid diagonal line represents ideal relationship in both graphs.

Figure 3.

This graph shows receiver operating characteristic curve for the ability of corrected adrenal gland volume (cAGV) based on 3D volume calculation ($\text{cAGV} > 420 \text{ mm}^3/\text{kg}$: Sensitivity 85%, specificity 95%, +LR 17.9, -LR 0.2) and cAGV based on 2D volume calculation ($\text{cAGV} > 433 \text{ mm}^3/\text{kg}$: Sensitivity 75%, specificity 89%, +LR 6.8, -LR 0.3) to predict a delivery within 7 days of scan.

Figure 4.

Bland-Altman plot for percentage of mean difference and 95% limits of agreement for inter-observer measurements performed by (a) 3D measurements of observer 1 and 2 (b) 2D measurements of observer 1 and 2.

Figure 5.

Bland-Altman plot for percentage of mean difference and 95% limits of agreement for intra-observer measurements performed by (a) 3D first and third measurements (b) 2D first and third measurements.

Table 1
Demographic and outcome characteristics (N=166)

Variables	PTL/PPROM (n=93)	Control (n=73)	p
Demographic characteristics			
Age (years, mean, SD) [†]	26.9 (6.07)	26.8 (6.32)	NS
Parity (median, range) [‡]	1 (0-4)	1 (0-9)	NS
Race (n, %) [§]			
Black	44 (47)	25 (34)	NS
Caucasian	35 (38)	24 (34)	
Hispanic	10 (11)	13 (18)	
Other	4 (4)	11 (14)	
GA at enrollment (weeks, mean, SD) [†]	28.3 (3.63)	28.2 (7.76)	NS
History of prior PTB (n, %) [§]	30 (33)	6 (8)	0.0002
Presence of PPROM (n, %) [§]	30 (32)	1 (1.4)	<0.0001
Outcome characteristics			
Delivery interval (days, median, range) [‡]	14.0 (0-112)	66.0 (1-171)	<0.0001
Gestational age at delivery (weeks, mean, SD) [†]	32.2 (4.63)	39.2 (2.10)	<0.0001
Delivery weight (grams, mean, SD) [†]	1842 (819.57)	3293 (537.24)	<0.0001
Delivery <37 weeks (n, %) [§]	77 (83)	6 (8)	<0.0001
Delivery within 7 days (n, %) [§]	38 (41)	2 (3)	<0.0001

SD: Standard deviation, PTL: preterm labor; PPROM: Preterm premature ruptures of membrane; GA: gestational age; PTB: preterm birth; SD: standard deviation; NS: non-significant

[†] Student t-test,

[‡] Mann-Whitney U test,

[§] Fisher's exact test

Table 2
Summary of precision results for inter-method, intra- and inter-observer variability

Statistic	Inter-method	Intra-observer		Inter-observer	
		3D volume	2D volume	3D volume	2D volume
Changes in mean (%)	1.3	0.7	-2.65	-0.5	-6.9
Standard error (%)	20.9	15.4	23.8	6.1	7.8
Coefficient of repeatability (%)	41.8	30.8	57.6	12.2	15.6
Interclass correlation coefficient (ICC)	0.979 (0.971-0.984)	0.982 (0.976-0.987)	0.925 (0.899-0.945)	0.998 (0.994-0.999)	0.997 (0.992-0.999)

Table 3
Predictive value of 3D and 2D derived corrected adrenal gland volumes for identification of women at risk for preterm birth

	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	-LR (95% CI)
3D volume > 420mm ³ /kg	85 (70-94)	95 (90-98)	17.9 (15.6-20.4)	0.2 (0.1-0.5)
2D volume > 433mm ³ /kg	75 (59-87)	89 (82-94)	6.8 (5.6-8.2)	0.3 (0.1-0.6)

+LR: positive likelihood ratio, -LR: negative likelihood ratio, CI: Confidence interval