Repeatability of volume calculations of the fetal urinary bladder

Trude Sætherhaug^{1§*}, Marit Hølmo Fasting^{1§}, Maria Aurora H. Røset^{1,2}, Mike Naylor³, Harm-Gerd K. Blaas^{1,2}, Torbjørn M. Eggebø^{1,2}

¹Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway ²National Center for Fetal Medicine, Trondheim University Hospital (St. Olavs Hospital), Trondheim, Norway ³Norwegian Center for Mathematics Education, Trondheim, Norway

*Corresponding author: Trude Sætherhaug, Email: trudeshaug@gmail.com

[§]These authors contributed equally to this work.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations used: VOCAL, virtual organ computer-aided analysis; SonoAVC, sonography-based automated volume count; ICC, intraclass correlation coefficient; CI, confidence interval; SD, standard deviation

Received December 1, 2015; Revision received April 4, 2016; Accepted April 9, 2016; Published May 22, 2016

ABSTRACT

OBJECTIVES: To investigate the intra- and inter-observer variation of the manual 3D Virtual Organ Computer-aided Analysis (VOCAL) in calculating fetal bladder volume and compare results from VOCAL with Sonography-based Automated Volume Count (SonoAVC) and a mathematical calculation based on three radiuses.

METHODS: Forty-five women attending the second trimester routine scan at St. Olavs Hospital, Trondheim, Norway were included in a prospective observational study. 3D acquisition of the fetal bladder was performed with Voluson E8 and the results were compared with intra-class correlation coefficient (ICC), repeatability coefficient and limits of agreement.

RESULTS: Good intra-observer agreement for both examiners was observed. ICCs were 0.98 and 0.99 and the repeatability coefficients were ± 27 and ± 25 mm³. However, we found a significant difference between the two observers. The mean difference was 20 mm³ (95% CI 9–32). Inter-observer ICC was 0.96. The mean volume based on radiuses was 148 (SD 14) mm³ vs. 145 (SD 14) mm³ based on VOCAL and 161 (SD 16) mm³ based on SonoAVC. The mean volume based on SonoAVC was significant larger than volumes based on radiuses. The ICCs between methods were good and varied from 0.82 to 0.93. The mean time used in calculating volumes was 94 seconds using VOCAL, 86 seconds using SonoAVC and 28 seconds calculating from radiuses.

CONCLUSIONS: VOCAL and SonoAVC did not show advantages compared to calculations based on radiuses. Therefore, we recommend calculating volumes of spherical fetal organs from three radiuses because this method was easiest to perform.

Keywords: 3D ultrasound, fetal bladder volume, repeatability analyses, volume calculations

INTRODUCTION

Evaluation of the fetal bladder is part of the routine fetal anatomy examination. Usually the bladder size is reported as small, normal or enlarged based on a subjective assessment. However, more precise measurement of bladder size is important in fetuses with anomalies involving the urinary tract, where the production of urine and the bladder filling are compromised. Such anomalies are renal failure or lower urinary tract obstruction [1-3]. A large urinary bladder is also associated with aneuploidy and Prune Belly syndrome [4-6]. In monitoring multiple pregnancies with twin-to-twin transfusion syndrome the difference in bladder size is important [7].

Three-dimensional (3D) ultrasound techniques allow assessment of organ volumes instead of two-dimensional (2D) measurements of diameters. Volumes can be calculated from manual geometric segmentations of the border of the organ [8-10]. With the help of automatic or semi-automatic algorithms, estimation of volumes is possible if the echogenicity of the organ differs significantly from the echogenicity of surrounding structures [10,11]. In spherical structures, volumes can be calculated mathematically from radiuses of the organ. Standardization and definitions in measurement methodology is warranted [12].

In this study we aimed to investigate the intra- and inter-observer variation of the manual 3D Virtual Organ Computer-aided Analysis (VOCAL) in calculating fetal bladder volume and compare results from VOCAL with Sonography-based Automated Volume Count (SonoAVC) and a mathematical calculation based on three radiuses.

METHODS

Forty-nine women attending the second trimester routine scan at the National Centre of Fetal Medicine, St. Olavs Hospital, Trondheim University Hospital, Norway from October 2014 to March 2015 participated. This study was a part of a study investigating ultrasound measurements of the fetal urinary tract. Women with singleton pregnancies without fetal malformations and gestational age between 17 and 20 weeks were eligible for the study. All women gave written consent and the local ethics committee approved the study (Rek Midt 2014/490).

All 3D acquisitions were performed by one expert in fetal medicine using Voluson E-8, (GE Medical Systems, Zipf, Austria), with a 3.5–7.5 MHz 3D curved multifrequency transabdominal transducer. The examiner tried to avoid shadows from fetal bony structures and stored three acquisitions from each fetus. The acquisition with best quality was selected for repeatability analyses. The selected acquisition was transferred to a personal computer and analyzed using 4D view software (GE Medical Systems, Zipf, Austria). Two non-experienced sonographers analyzed the 3D-volumes. They trained by calculating 20 cases in each method before starting the study.



Figure 1. Calculation based on VOCAL.



Figure 2. Calculation based on SonoAVC.

Three different methods of estimating volumes were compared. The first method used was the VOCAL method illustrated in **Figure 1**. VOCAL allows the user to manually define the volume of interest by tracing around the border in a number of steps as the dataset is rotated 180 degrees, and 15 degrees steps were used in this study. Two examiners performed three volume assessments from each fetus and the results were blinded on the computer screen and stored on the hard disk. The next method used was SonoAVC (**Fig. 2**). SonoAVC is based on the

difference in echogenicity between the organ of interest, and the tissue surrounding it. Region of interest is adjusted, then SonoAVC identifies the volume by giving the structure examined a specific color, and provides automated measurements of its volume [11]. At last the volumes were calculated from measurements of the greatest diameter in three different sectional planes using the formula $4/3\pi$ radius ($r_1 * r_2 * r_3$) (**Fig. 3**). The different methods were used with several days interval to avoid bias.

The mean time spent in calculating volumes from stored 3D volumes were assessed in five acquisitions. The time used in adjusting region of interest and thresholds was included when assessing SonoAVC.

Normality plots were used to assess the distribution of measurements obtained. Intra-observer repeatability of the measurements was expressed as the difference between the highest and lowest measurements, the intra-class correlation coefficient (ICC) and the repeatability coefficient. The repeatability coefficient defines the range within which two measurements by the same observer will fall for 95% of subjects. The differences between the first, second and third measurements were evaluated with three-way analysis of variance.

The analysis of inter-observer agreement was performed using the mean of the three measurements from each of the examiners. If zero was inside the 95% CI of the difference, no bias was assumed. To assess systematic bias between the examiners, differences between values were plotted against means of the measurements. Limits of agreement with 95% CIs of the lower and upper limits were calculated as described by Bland and Altman [13]. Inter-observer agreement was also expressed using the two-way random effects ICC. The analysis of inter-method agreement was performed using the same methods as described in inter-observer variation. Statistical analyses were performed with IBM SPSS Statistics for Mac, v. 21.0 Armonk, NY: IBM Corp.

Table 1. Intra-observer repeatability for volume measurements of fetal urinary bladders (VOCAL).

				_		Difference between highest and lowest values				st values
Measure- ment	Mean	Median	Range	ICC (95% CI)	Repeatability coefficient	Mean	Median	10th centile	90th centile	Range
Examiner 1	155	151	(29–384)	0.98 (0.97–0.99)	± 27	18	14	5	33	(3–70)
Examiner 2	135	125	(5–432)	0.99 (0.99–1.00)	± 25	15	11	3	31	(0–61)

Mean, median and range of measurements are calculated from the mean of three measurements; CI, confidence interval; ICC, intraclass correlation coefficient

Table 2. Inter-observer repeatability in volume calculation of fetal urinary bladders (VOCAL).

					Difference between the two observers				
	Mean	Median	Range	ICC (95% CI)	Mean (95% CI)	1.96 SD	Lower limit (95% Cl)	Upper limit (95% Cl)	Range
Inter-observer agreement	145	134	(17–408)	0.96 (0.93–0.98)	20 (9–32)	74	-54 (-7335)	94 (75–113)	(-68–160)

Mean, median and range for volume are calculated from the mean results from the two examiners; CI, confidence interval; ICC, intraclass correlation coefficient; SD, standard deviation

RESULTS

The size and volume of the bladder were successfully visualized in 45 of the 49 women, but could not be analyzed in four of cases due to shadowing from fetal bone structures. Median maternal age was 32 years (range 22 to 42) and 42 % were nulliparous women. Median gestational age was 18 weeks + 4 days (range 17 to 20 weeks).

The results of the analysis of intra-observer repeatability using the VOCAL method of the two observers are presented in **Table 1**. The mean values for the first, second and third measurements were 153 (SD 88), 155 (SD 89) and 156 (SD 88) mm³, respectively for the first examiner, and 132 (SD 103), 135 (SD 102) and 136 (SD 102) mm³, respectively, for the second examiner. There was no systematic differences between first, second or third measurements (P = 0.52 and P = 0.15) for any of the examiners, but a tendency to greater variation associated to large

volumes for both examiners.

The inter-observer variation is presented in **Table 2**. The mean difference was 20 mm³ (95% CI 9–32). There was no systematic variation over the range of measured values, however, the CI is not crossing zero, which indicates a systematic difference between the two observers. Limits of agreement are shown in **Figure 4A** and in **Table 2**.

The mean volumes calculated using VOCAL (mean VOCAL values from the two observers) was 145 (SD 94) mm³, using SonoAVC 161 (SD 107) mm³ and using radiuses 148 (SD 92) mm³. The results of the analyses of inter-method agreement between VOCAL and SonoAVC are presented in **Table 3**. The mean difference was -16 mm³ (95% CI -27 to -5 mm³). The CI is not crossing zero indicating a systematic difference between methods. No systematic variation over the range of measured values was observed (**Fig. 4B**). Limits of agreement are reported in **Table 3**.



Figure 3. Calculation based on radiuses ((4/3) $\pi^*r_1^*r_2^*r_3$).



Figure 4. Bland-Altman plots with mean difference in mm³ (____), 95 % confidence interval (....) and 95 % limits of agreement (*i.e.*, mean difference ± 1.96 SD) (----) shown. A. For inter-observer agreement. B. For inter-method agreement between VOCAL and SonoAVC. C. For inter-method agreement between VOCAL and radiuses. D. For inter-method agreement between SonoAVC and radiuses.

The results of the analyses of inter-method agreement between VO-CAL and the volume calculation from radiuses are presented in **Table 3**. The mean difference was minus 3 mm³ (95% CI -16 to 10 mm³). The CI is including zero indicating no systematic difference between methods. We observed a tendency to greater variation of large values (**Fig. 4C**). Limits of agreement are reported in **Table 3**.

The results of the analyses of inter-method agreement between SonoAVC and volume calculation from radiuses are presented in **Table** **3**. The mean difference was 13 mm³ (95% CI -5 to 31 mm³). The CI is including zero indicating no systematic difference between the methods. We observed a tendency to greater variation of large values (**Fig. 4D**). Limits of agreement are reported in **Table 3**.

The mean time used in calculating the 3D volumes was 94 seconds using VOCAL, 86 seconds using SonoAVC and 28 seconds calculating from radiuses.

Table 5. Inter-method agreement between VOCAL, SONDAVC and radiuses in calculations of retai bladder volume	Table 3.	Inter-method agreement	t between VOCAL	, SonoAVC and radiuses in	calculations of feta	I bladder volume.
---	----------	------------------------	-----------------	---------------------------	----------------------	-------------------

					Difference between the two methods				
	Mean	Median	Range	ICC (95% CI)	Mean (95% CI)	1.96 SD	Lower limit (95% Cl)	Upper limit (95% CI)	Range
Vocal vs. SonaAVC	153	148	(14–424)	0.93 (0.88–0.96)	-16 (-275)	73	-89 (-108–-70)	57 (38–76)	(-146-40)
Vocal vs calculations from radiuses	146	143	(12–442)	0.89 (0.81–0.94)	-3 (-16–10)	86	-89 (-111–-67)	83 (61–105)	(-80–102)
SonoAVC vs. calculation from radiuses	154	141	(9–459)	0.82 (0.69–0.90)	13 (-5–31)	118	-104 (-13474)	130 (100–160)	(-68–212)

Mean, median and range for volume are calculated from the mean of the two methods; CI, confidence interval; ICC, intraclass correlation coefficient; SD, standard deviation; VOCAL, Virtual Organ Computer-aided Analysis; SonoAVC, Sonography-based Automated Volume Count

DISCUSSION

We found very good intra-observer repeatability for both examiners using the VOCAL method, but observed a systematic variation between the two examiners. Volumes based on SonoAVC were slightly larger than volumes based on VOCAL. The inter-method agreement between VOCAL and mathematical calculation based on radiuses was good.

The size of an organ is better expressed as a volume, than from diameters in two dimensions. In the first trimester, volumes of the gestational sac [14], the trophoblast [15], the placenta [16] and the embryo and embryonic structures can be calculated [17,18]. In the second and third trimester organ volumes are of interest, however, the clinical value of calculating volumes is not documented. It has been suggested to use volume calculations of fetal bladder in monochoriote twin pregnancies [19]. These fetuses have a 10% risk of developing twin-to-twin-transfusion syndrome. Difference in bladder size between the fetuses is an important diagnostic criterion for laser coagulation of placental vessels [20,21]. Complete fetal urethral obstruction is a lethal anomaly [22]. The renal function is destroyed following extreme hydronephrosis causing oligohydramnios and insufficient development of the lungs. The enlarged fetal bladder, often called fetal megacystis, is easy to observe by ultrasound in the first trimester and usually defined from measuring one diameter. Obstruction of bladder outflow is usually caused by urethral valves or atresia, but can also be a complication of an obstructive uretererocoele [4]. An ureterocele might develop in fetuses with double ureteres, and the ureterocele can obstruct bladder outflow. Fetal megacystis or enlarged urinary bladder is defined as greatest diameter > 7 mm in the first trimester [6,23], but the definition varies in the second and third trimester. One definition includes an enlarged bladder with failure to empty over a period of forty-five minutes [4]. Volume calculation might be recommended in considering the prognosis of fetuses with enlarged bladder. Especially might changes in bladder volume over time be of interest differentiating between complete or partial obstructions. If a fetal ureterocele is diagnosed, close surveillance of the fetus is recommended [24], and calculating bladder volume might be helpful. Further studies are needed to evaluate clinical benefits.

VOCAL has shown a high degree of accuracy in measuring the volume of irregular objects *in vitro* [25]. However, the accuracy will depend on the number of steps used rotating the organ. We observed a significant inter-observer variation using the VOCAL method. The examiners were thoroughly explained to draw the lines exactly around the border of the fetal bladder. Still a subjective variance might occur whether the line is drawn on the echogenic part or on the inside of the demarcation line. **Figure 1** illustrates that the exact border of the urinary bladder is slightly unclear and that volume calculations will depend on subjective variations. Ultrasound artifacts can also make the demarcation line unclear, especially at the lateral borders where the ultrasound beams run parallel to the border of the organ.

SonoAVC is a tool for automatic volume calculations. The calculation is easy to perform from a 3D acquisition and has shown high accuracy and better validity than VOCAL in measuring follicular volumes in women [11]. Different echogenicity and clear demarcation between examined organs are a prerequisite using SonoAVC. The threshold has to be adjusted on the ultrasound device using the automatic SonoAVC calculation and the examiner should observe that the calculation seems appropriate. Calculation based on radiuses is quicker to perform than with VOCAL and SonoAVC. The main limitation of the mathematical volume from radiuses is that only spherical volumes can be calculated.



The fetal bladder usually has a spherical shape. Other fetal organs like the stomach, renal pelvises and the brain ventricles have an irregular shape and VOCAL or SonoAVC should be preferred methods analyzing volumes of these organs [10].

Softwares for volume calculations are implemented in many modern ultrasound devices. It is relatively easy to learn how to use the methods, but the 3D techniques were time consuming, as also shown in other studies [10,25]. The time needed using VOCAL depends on the number of steps used in the calculation [25]. The cost of including VOCAL and SonoAVC into the ultrasound device was around 3500 US dollars each.

Strengths of this study are that the results on the computer screen were blinded when the examiners calculated the volumes and that interval of several days were used estimating volumes with different methods. A standard pre-set of the ultrasound device and the computer were used to avoid bias. Limitations are related to the size of the study population and that four volumes could not be analysed due to shadowing from bony structures in the fetus. Factors complicating fetal volume assessments are absence of clearly defined anatomical landmarks for measurement and the lacking possibility in validating fetal measurements in vivo against a reference standard [12]. As the true bladder volumes are not known, we could only estimate differences between examiners and methods. Berg et al. compared manually segmented ultrasound volume reconstructions with known in-vitro volumes and found that the reconstructions were accurate and repeatable [17]. Raine-Fenning et al. compared VOCAL assessments with true volumes in-vitro, they found that the ultrasound method overestimated volumes, but the overestimation was less than five percent. VOCAL was less precise using 30 degrees rotational steps, but no significant differences in overall validity between the 15°, 9° and 6° rotation steps were observed [25]. Postnatal studies might compare sonographically assessed bladder volumes with true volumes in infants needing catheterization on clinical indication, but we will never know the exact fetal bladder volume in utero.

In conclusion, VOCAL and SonoAVC did not show advantages compared to calculations based on radiuses. The philosophical principle of William of Ockham says: "One should not make more assumptions than the minimum needed such that we should choose from a set of otherwise equivalent models the simplest one". Therefore, we recommend calculating volumes of spherical fetal organs from three radiuses because this method is easiest to perform.

Acknowledgments

None

Clinical trial registration

All scans were done during the routine second trimester examination. The study was approved by the Norwegian ethical committee (Rek Midt 2014/490).

References

- Bondagji NS (2014) Antenatal diagnosis, prevalence and outcome of congenital anomalies of the kidney and urinary tract in Saudi Arabia. Urol Ann 6: 36-40. doi: 10.4103/0974-7796.127021. PMID: 24669120
- Eckoldt F, Heling KS, Woderich R, Wolke S (2004) Posterior urethral valves: prenatal diagnostic signs and outcome. Urol Int 73: 296-301. doi: 10.1159/000081586. PMID: 15604570
- 3. McHugo J, Whittle M (2001) Enlarged fetal bladders: aetiology, management

and outcome. Prenat Diagn 21: 958-963. PMID: 11746149

- Fievet L, Faure A, Coze S, Harper L, Panait N, et al. (2014) Fetal megacystis: etiologies, management, and outcome according to the trimester. Urology 84: 185-190. doi: 10.1016/j.urology.2014.02.018. PMID: 24745800
- Kraiem J, Ben Brahim Y, Jouini H, Ben Brahim I, Chatti S, et al. (2006) [Prune belly syndrome: early prenatal diagnosis and management]. Tunis Med 84: 458-461. PMID: 17039740
- Liao AW, Sebire NJ, Geerts L, Cicero S, Nicolaides KH (2003) Megacystis at 10-14 weeks of gestation: chromosomal defects and outcome according to bladder length. Ultrasound Obstet Gynecol 21: 338-341. doi: 10.1002/uog.81. PMID: 12704740
- Wee LY, Fisk NM (2002) The twin-twin transfusion syndrome. Semin Neonatol 7: 187-202. PMID: 12234743
- Blaas HG, Eik-Nes SH, Berg S, Torp H (1998) In-vivo three-dimensional ultrasound reconstructions of embryos and early fetuses. Lancet 352: 1182-1186. doi: 10.1016/S0140-6736(98)03227-9. PMID: 9777835
- Duin LK, Willekes C, Vossen M, Beckers M, Offermans J, et al. (2008) Reproducibility of fetal renal pelvis volume measurement using three-dimensional ultrasound. Ultrasound Obstet Gynecol 31: 657-661. doi: 10.1002/uog.5208. PMID: 18470970
- Kusanovic JP, Nien JK, Gonçalves LF, Espinoza J, Lee W, et al. (2008) The use of inversion mode and 3D manual segmentation in volume measurement of fetal fluid-filled structures: comparison with Virtual Organ Computer-aided AnaLysis (VOCAL). Ultrasound Obstet Gynecol 31: 177-186. doi: 10.1002/ uog.5242. PMID: 18254130
- Raine-Fenning N, Jayaprakasan K, Clewes J, Joergner I, Bonaki SD, et al. (2008) SonoAVC: a novel method of automatic volume calculation. Ultrasound Obstet Gynecol 31: 691-696. doi: 10.1002/uog.5359. PMID: 18484679
- Ioannou C, Sarris I, Salomon LJ, Papageorghiou AT (2011) A review of fetal volumetry: the need for standardization and definitions in measurement methodology. Ultrasound Obstet Gynecol 38: 613-619. doi: 10.1002/uog.9074. PMID: 21674657
- Bland JM, Altman DG (2003) Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol 22: 85-93. doi: 10.1002/ uog.122. PMID: 12858311
- Odeh M, Hirsh Y, Degani S, Grinin V, Ofir E, et al. (2008) Three-dimensional sonographic volumetry of the gestational sac and the amniotic sac in the first trimester. J Ultrasound Med 27: 373-378. PMID: 18314515
- Reus AD, El-Harbachi H, Rousian M, Willemsen SP, Steegers-Theunissen RPM, et al. (2013) Early first-trimester trophoblast volume in pregnancies that result in live birth or miscarriage. Ultrasound Obstet Gynecol 42: 577-584. doi: 10.1002/ uog.13197. PMID: 23996572
- 16. Cheong KB, Leung KY, Li TKT, Chan HY, Lee YP, et al. (2010) Comparison of inter- and intraobserver agreement and reliability between three different types of placental volume measurement technique (XI VOCAL, VOCAL and multiplanar) and validity in the in-vitro setting. Ultrasound Obstet Gynecol 36: 210-217. doi: 10.1002/uog.7609. PMID: 20201116
- Berg S, Torp H, Blaas HG (2000) Accuracy of in-vitro volume estimation of small structures using three-dimensional ultrasound. Ultrasound Med Biol 26: 425-432. PMID: 10773373
- Blaas HG, Taipale P, Torp H, Eik-Nes SH (2006) Three-dimensional ultrasound volume calculations of human embryos and young fetuses: a study on the volumetry of compound structures and its reproducibility. Ultrasound Obstet Gynecol 27: 640-646. doi: 10.1002/uog.2794. PMID: 16710882
- Kist WJ, Slaghekke F, Papanna R, Johnson A, Vandenbussche FP, et al. (2011) Sonography-based Automated Volume Count to estimate fetal urine production in twin-to-twin transfusion syndrome: comparison with Virtual Organ Computeraided AnaLysis. Am J Obstet Gynecol 205: 574. doi: 10.1016/j.ajog.2011.06.090. PMID: 21906713
- Sueters M, Oepkes D (2014) Diagnosis of twin-to-twin transfusion syndrome, selective fetal growth restriction, twin anaemia-polycythaemia sequence, and twin reversed arterial perfusion sequence. Best Pract Res Clin Obstet Gynaecol 28: 215-226. doi: 10.1016/j.bpobgyn.2013.12.002. PMID: 24433823
- Diehl W, Diemert A, Hecher K (2013) Twin-twin transfusion syndrome: treatment and outcome. Best Pract Res Clin Obstet Gynaecol 28: 227-238. doi: 10.1016/j. bpobgyn.2013.12.001. PMID: 24512749
- Morris RK, Middleton LJ, Malin GL, Quinlan-Jones E, Daniels J, et al. (2015) Outcome in fetal lower urinary tract obstruction: a prospective registry study. Ultrasound Obstet Gynecol 46: 424-431. doi: 10.1002/uog.14808. PMID: 25689128
- 23. Sepulveda W (2004) Megacystis in the first trimester. Prenat Diagn 24: 144-149.

doi: 10.1002/pd.790. PMID: 14974124

- 24. Godinho AB, Nunes C, Janeiro M, Carvalho R, Melo MA, et al. (2013) Ureterocele: antenatal diagnosis and management. Fetal Diagn Ther 34: 188-191. doi: 10.1159/000353388. PMID: 23969704
- 25. Raine-Fenning NJ, Clewes JS, Kendall NR, Bunkheila AK, Campbell BK, et al. (2003) The interobserver reliability and validity of volume calculation from

three-dimensional ultrasound datasets in the *in vitro* setting. Ultrasound Obstet Gynecol 21: 283-291. doi: 10.1002/uog.61. PMID: 12666225



This work is licensed under a Creative Commons Attribution-Non-Commercial-ShareAlike 4.0 International License: http://creativecommons.org/licenses/by-nc-sa/4.0

