

Comparison of Manual and Automated Measurements of Monodominant Follicle Diameter with Different Follicle Size in Infertile Patients

Ping Pan^{1,2}, Xiaoli Chen¹, Yu Li¹, Qingxue Zhang¹, Xiaomiao Zhao¹, Madafeitom M. A. Bodombossou-Djobo^{1,3}, Dongzi Yang^{*}

1 Department of Obstetrics and Gynecology, Memorial Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, People's Republic of China, **2** Department of Reproductive Medicine, Guangdong Provincial Family Planning Research Institute, Guangzhou, Guangdong, People's Republic of China, **3** Fertility Medical Center, Guangzhou Elizabeth Women's Hospital, Guangzhou, Guangdong, People's Republic of China

Abstract

This study evaluated the consistency of manual and automated measurements of monodominant follicle diameter with different follicle size in infertile patients. Transvaginal two-dimensional (2D) ultrasound and SonoAVC (Sonography-based Automated Volume Calculation) were both performed in 226 infertile patients with monodominant follicle growth. 2D diameters were separately compared with SonoAVC-generated d(V) and m-d values in different follicle category, i.e. >10 to 14 mm, >14 to 18 mm, >18 to 22 mm and >22 mm. There was moderate degree of consistency between 2D diameter and SonoAVC-generated parameters regardless of follicle size. The mean differences were 0.82 mm between 2D diameter and SonoAVC-generated d(V) value, and 0.22 mm between 2D diameter and SonoAVC-generated m-d value, respectively. The discrepancy of manual and automated measurements tended to increase as follicle size increased. Our study suggested that compared with manual measurement, SonoAVC might underestimate follicle size. The absolute size of a follicle affected the consistency of two techniques.

Citation: Pan P, Chen X, Li Y, Zhang Q, Zhao X, et al. (2013) Comparison of Manual and Automated Measurements of Monodominant Follicle Diameter with Different Follicle Size in Infertile Patients. PLoS ONE 8(10): e77095. doi:10.1371/journal.pone.0077095

Editor: Sanjoy Bhattacharya, Bascom Palmer Eye Institute, University of Miami School of Medicine, United States of America

Received: May 27, 2013; **Accepted:** August 29, 2013; **Published:** October 10, 2013

Copyright: © 2013 Pan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Funding for this work was provided by: Sun Yat-Sen University Clinical Research 5010 Program (Grant No. 2007-017); The Science Technology Research Project of Guangdong Province (2012A030400010); The Science Research Project of Population and Family Planning Committee of Guangdong Province (Grant No. 2012316). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: yangdz@mail.sysu.edu.cn

Introduction

To date, transvaginal two-dimensional (2D) ultrasound has been widely used to track ovarian follicle growth, confirm ovulation and detect ovulatory dysfunction. 2D follicular diameter is a common ultrasonic parameter for the assessment of follicle size, which is conventionally restricted to manual determination. More recently, the advent of SonoAVC (Sonography-based Automated Volume Calculation) has allowed us to assess follicle size automatically. SonoAVC can identify and quantify hypoechoic regions within a three-dimensional (3D) dataset, such as a follicle in an acquired ovarian volume, and provide automatic estimation of their dimensions [1].

Previous studies have demonstrated that follicular volume measured by SonoAVC is more accurate than that estimated from 2D ultrasound, and equivalent to the aspirated follicular fluid volume [2-5]. Besides follicular volume, SonoAVC as well

as provides other ultrasonic parameters, such as the volume-based diameter d(V) and mean follicular diameter (m-d). It has been shown that SonoAVC-generated d(V) and m-d values correlated extremely well with the manual measurements by 2D ultrasound, and the observed differences were <1 mm. Yet, some studies used SonoAVC to determine multiple growing follicles in hyperstimulated ovaries, and evaluated follicles ranging in wide diameter from 2.3 to 32 mm, some studies did not list the particular size of the follicles [6,7].

SonoAVC provides the opportunity to apply volumetric measure to track follicle growth and assess follicle size automatically, however, the cutoff value of follicular volume as a new parameter needs to be established. As a result, the use of SonoAVC-generated d(V) and m-d values but not follicular volume as surrogates of 2D diameter is of clinical relevance in practice at present time. Moreover, the applicability of SonoAVC still requires further study, and it is not clear whether there would be a difference of the consistency of manual and

automated measurements for dominant follicles with various size. To clarify this issue, in the present investigation we compared SonoAVC with 2D ultrasound for the assessment of a cohort of dominant follicles in various size category. Meanwhile, to avoid possible confounding effects of multiple developing follicles on the reliability of our study, we preferred to use the experimental model of monodominant follicle growth in infertile patients.

Materials and Methods

Subjects

The infertile patients undergoing infertility treatment in outpatient clinic, or artificial insemination and frozen-thawed embryo transfer in the reproductive centre of Memorial Hospital of Sun Yat-Sen University from July 2012 to November 2012 were recruited. Only natural and mildly stimulated cycles with pre-ovulatory monodominant follicle growth (>10 mm in diameter) were included in this study, and the cycles with two or more developing follicles >10 mm were excluded. The protocols of mildly ovarian stimulation were following: (a) clomiphene citrate with a dosage of 50 mg/day from day 3 to day 7; (b) letrozole with a dosage of 5 mg/day from day 3 to day 7. A total of 226 infertile patients contributed to this study. The study was approved by the Human Research and Ethics Committee of Sun Yat-Sen Memorial Hospital, and all patients gave written informed consent.

2D follicular diameter measurement

A transvaginal scan was performed using a General Electric Voluson E8 Expert instrument (GE Medical Systems, Zipf, Austria) equipped with a 5 to 9 MHz transvaginal microconvex volume probe. All ultrasound examinations were performed by a single experienced investigator (Ping Pan) to avoid interobserver variation. By using the real-time 2D mode, two orthogonal diameters (d1 and d2) at the largest follicle plane were determined by placing calipers at the inner follicle border, as the standard clinical practice in our reproductive centre. The vaginal probe was placed to the follicle as close as possible, in order to make follicle borders visualized clearly. Each 2D follicular diameter was examined three times consecutively during the examination. Mean follicular diameter in each measurement corresponded to $(d1 + d2)/2$, then the average value from three measurements for each follicle was used for statistical analysis.

SonoAVC scan

After completion of 2D measurement, SonoAVC was initiated to capture a 3D volume of each follicle, as reported in the literatures [1,7,8]. In brief, the sweep angle in 3D volume mode was set to 120° so as to include the entire ovary, and a 3D dataset was acquired using the high-quality, slow-sweep mode. Similarly, SonoAVC scan was also replicated for three times for each follicle. Post-processing of SonoAVC was performed with 4D View (version 10.3; GE Medical Systems) by the same investigator (Ping Pan) at least one month after data collection to prevent the recall of 2D examination. The settings of growth

and separation within the software were kept uniform at default values of 'mid' for all follicle measurements.

The ultrasonic parameters generated by SonoAVC included d(V), dx, dy, dz, m-d and V, but according to the study's design, we only enrolled d(V), m-d and V for analysis. In detail, d(V) was the diameter of a perfect sphere with the same volume as the follicle, and m-d was the arithmetic mean of the three longest orthogonal diameters (dx, dy and dz), and V was the volume of the identified follicle based on the voxel count within it [7]. Accordingly the average values of all SonoAVC-generated parameters from three repeated measurements for each follicle were analyzed.

Statistical Analysis

Statistical analysis was performed by SPSS software (Statistical Package for Social Sciences, SPSS Inc, Chicago, IL, USA) version 13.0. Mean and SD were used for description of variables. Since it is hypothesized that the absolute size of a follicle might affect the consistency of manual and automated measurements, subgroup analysis was conducted based on the absolute follicular size. Consequently, all pre-ovulatory monodominant follicles were arbitrarily sorted into four groups based upon 2D follicular diameters, i.e. >10 to 14mm, >14 to 18mm, >18 to 22mm and >22mm, as these thresholds had clinical importance according to previous studies [7,9].

The intra-observer reliability of 2D measurement and SonoAVC was assessed by intraclass correlation coefficients (ICCs) and its 95% confidence intervals (CI) using a one-way random model, by calculating those 2D and SonoAVC-generated parameters obtained from three measurements for each subject. Correlation between 2D follicular diameter and SonoAVC-generated d(V), m-d and V measurements were separately evaluated using the Pearson correlation coefficient. The concordance for SonoAVC and manual 2D measurement was calculated by ICCs with absolute agreement and 95% CI using a two-way mixed model. The ICCs range from 0 to 1, and values under 0.10 are considered virtually none, 0.11-0.40 slight, 0.41-0.60 fair, 0.61-0.80 moderate, and 0.81-1.00 substantial [10]. Also, we used the limits of agreement method as reported by Bland and Altman [11]. The mean difference, the upper and lower limits of agreement were calculated and presented. The range between the upper and lower limits of agreement provides more information to judge how well two methods of measurement agree. The smaller the range between the upper and lower limits of agreement the better the agreement is [12]. P values of <.05 were considered significant.

Results

General characteristics

The mean age of participants was 32.21±4.25 years (range 23-44 years). Natural cycles accounted for 53.10% (120/226), and mildly stimulated cycles accounted for 46.90% (106/226). Of the participants, 53.10% (120/226) performed artificial insemination, 37.61% (85/226) performed frozen-thawed embryo transfer, and 9.29% (21/226) underwent infertility treatment in outpatient clinic. Two hundred and twenty-six pre-ovulatory monodominant follicles were studied in 226 patients,

Table 1. Correlation coefficients of 2D follicular diameter and SonoAVC-generated parameters in various follicle category.

Ultrasonic parameters	Total (n=312)	Follicle size (mm)			
		>10 to 14 (n=46)	>14 to 18 (n=107)	>18 to 22 (n=118)	>22 (n=41)
2D diameter versus SonoAVC d(V)	0.960	0.775	0.788	0.740	0.743
2D diameter versus SonoAVC m-d	0.958	0.738	0.775	0.723	0.718
2D diameter versus SonoAVC V	0.929	0.744	0.773	0.745	0.762
SonoAVC d(V) versus SonoAVC m-d	0.995	0.974	0.980	0.971	0.974

doi: 10.1371/journal.pone.0077095.t001

with 57.08% (129/226) from right ovary, and 42.92% (97/226) from left ovary. The number of ultrasonography per follicle in the menstrual cycle was 1 in 157 follicles, 2 in 52 follicles and 3 in 17 follicles, therefore 312 follicular data were enrolled in the analysis. Mean 2D follicular diameters were 18.09±3.37 mm (range 10.50-27.33mm), mean SonoAVC-generated d(V) values were 17.26±3.41 mm (range 9.80-28.13 mm), mean SonoAVC-generated m-d values were 17.87±3.51 mm (range 10.43-28.50 mm), mean SonoAVC-generated follicular volumes were 3.01±1.75 mL (range 0.49-11.65 mL).

Intraobserver reliability of 2D measurement and SonoAVC

The ICCs were 0.986 (95% CI: 0.983-0.989) for 2D follicular diameter, 0.992 (95% CI: 0.990-0.993) for SonoAVC-generated d(V) value, and 0.993 (95% CI: 0.991-0.994) for SonoAVC-generated m-d value, suggesting good intraobserver reliability for each measurement.

Correlation coefficient of 2D measurement and SonoAVC

On the whole, there were strong positive correlations between 2D follicular diameter and SonoAVC-generated parameters. The Pearson correlation coefficient was 0.960 for 2D diameter and SonoAVC-generated d(V) value, 0.958 for 2D diameter and SonoAVC-generated m-d value, and 0.929 for 2D diameter and SonoAVC-generated V value, respectively (P=0.000). When manual measurement and SonoAVC were separately compared in each follicle category, we found a moderate correlation between two techniques, with Pearson correlation coefficient ranging from 0.718 to 0.788. On the contrary, correlation coefficient between SonoAVC-generated d(V) and m-d value was still high, as presented in Table 1.

The ICCs for 2D measurement and SonoAVC

Accordingly the ICCs were 0.933 (95% CI: 0.710-0.972) for 2D diameter and SonoAVC-generated d(V) value, and 0.955 (95% CI: 0.943-0.965) for 2D diameter and SonoAVC-

Table 2. The ICCs of 2D measurement and SonoAVC in various follicle category.

Ultrasonic parameters	Total (n=312)	Follicle size (mm)			
		>10 to 14 (n=46)	>14 to 18 (n=107)	>18 to 22 (n=118)	>22 (n=41)
2D diameter versus SonoAVC d(V)	0.933	0.602	0.665	0.587	0.585
95% confidence intervals	0.710-0.972	0.001-0.834	0.213-0.836	0.136-0.786	0.107-0.805
2D diameter versus SonoAVC m-d	0.955	0.711	0.742	0.679	0.687
95% confidence intervals	0.943-0.965	0.519-0.832	0.642-0.817	0.567-0.766	0.487-0.819

*. ICCs = intraclass correlation coefficients

doi: 10.1371/journal.pone.0077095.t002

generated m-d value (P=0.000). Table 2 illustrated the ICCs of manual 2D measurement and SonoAVC in each follicle category, and we found moderate concordance between two techniques, with ICCs ranging from 0.585 to 0.742. Regardless of follicle size, 2D diameter versus SonoAVC-generated d(V) value corresponded to a lower ICCs value when compared with 2D diameter versus SonoAVC-generated m-d value.

The limits of agreement between 2D measurement and SonoAVC

The mean differences were 0.82 mm between 2D diameter and SonoAVC-generated d(V) value, and 0.22 mm between 2D diameter and SonoAVC-generated m-d value, respectively. However, the observed differences in each follicle category increased as the follicle size increased, consequently the range of the limits of agreements were gradually wider in large follicle size as compared with in small follicle size. Unexpectedly, the observed differences in each follicle category between 2D diameter and SonoAVC-generated m-d value did not change substantially, but the range of the limits of agreements likewise appeared to be wider for large follicle size in comparison with small follicle size, shown in Table 3.

Discussion

Whenever a novel technique comes out, comparison between the new and conventional method is indispensable. This study was unique in that we not only used different

Table 3. Bland-Altman plot analysis of 2D measurement and SonoAVC in various follicle category.

Ultrasonic parameters	Total (n=312)	Follicle size (mm)			
		>10 to 14 (n=46)	>14 to 18 (n=107)	>18 to 22 (n=118)	>22 (n=41)
2D diameter versus SonoAVC d(V)					
Mean difference	0.82±0.95	0.68±0.62	0.77±0.92	0.86±1.00	1.02±1.19
Upper limit of agreement	-1.08	1.92	2.61	2.86	3.40
Lower limit of agreement	2.72	-0.56	-1.07	-1.14	-1.36
2D diameter versus SonoAVC m-d					
Mean difference	0.22±1.01	0.24±0.66	0.20±1.00	0.21±1.07	0.26±1.18
Upper limit of agreement	-1.80	1.56	2.20	2.35	2.62
Lower limit of agreement	2.24	-1.08	-1.80	-1.93	-2.10

doi: 10.1371/journal.pone.0077095.t003

statistical tools to evaluate the concordance of 2D ultrasound and SonoAVC, but also compared two techniques in various follicle size category for further detail. Our data showed that there was moderate degree of consistency between 2D follicular diameter and SonoAVC-generated d(V) and m-d measurements regardless of follicle size, either analyzed by Pearson correlation coefficient or ICCs. The mean differences between 2D diameter and SonoAVC-generated parameters were <1 mm, and interestingly the discrepancy of two techniques tended to increase with the follicle size, revealed by the limits of agreement.

The greatest strength of SonoAVC lies in accurate measurement of follicular volume, as mentioned above [2-5]. It is well known that a follicle is after all a three-dimensional structure, many investigators have already confirmed follicular volume is the most accurate measure of its size [2,3,7,13]. Because SonoAVC assessed follicular volume on the basis of the voxel count within it, this measurement seemed objective and reproducible [5]. Thus, SonoAVC had been considered to standardize the process of follicular assessment. What is more, even though SonoAVC required postprocessing, it decreased examination time both for patients and ultrasonographer, which helped to improve work flow in clinical practice [3,14].

Despite the aforementioned advantages of SonoAVC, it does take time for clinicians to shift from 2D follicular diameter to SonoAVC-generated follicular volume for follicle monitoring. For the moment the use of SonoAVC-generated d(V) and m-d values as surrogates of 2D diameter might be clinically relevant and practical, and it was the reason why we conducted this study. Moreover, we compared 2D ultrasound and SonoAVC

only for the assessment of monodominant follicle in the study, because in general monodominant follicle was round or ellipsoid, but in a hyperstimulated ovary with multifollicular growth, most follicles had irregular shape due to compression by adjacent follicles. We deemed that under such an ideal situation of monofollicular growth, estimation of the consistency between two techniques might be objective and accurate.

As a whole, Pearson correlation coefficient and ICCs calculated for manual 2D measurements and SonoAVC-generated d(V) and m-d values were both very high, predicting good concordance. This was in support with previous study [6]. When two techniques were compared in each follicle category, the results were somewhat different, Pearson correlation coefficient and ICCs lowered down regardless of follicle size, predicting moderate concordance. However, it was worthy to note that SonoAVC-generated m-d value was more comparable to 2D diameter in comparison with SonoAVC-generated d(V) value, which was evidenced by great ICCs of SonoAVC-generated m-d value and 2D diameter.

As far as we know, the Bland-Altman analysis was extensively used to evaluate agreement between two measurement methods in numerous published studies, herein we used the Bland-Altman method for detailed analysis. In our study, the mean differences between manual 2D measurement and SonoAVC were <1 mm, and SonoAVC-generated d(V) and m-d values tended to be smaller than 2D diameter, i.e. d(V) value was less than 2D diameter by 0.82 mm on average, and m-d value was less than 2D diameter by 0.22 mm on average. In other words, SonoAVC-generated m-d value was more identical to 2D diameter than SonoAVC-generated d(V) value, which was as same as the result analyzed by ICCs. Our results were a little distinct from a study by Ata et al., they observed the SonoAVC-generated d(V) and m-d measurements were both in good agreement with conventional 2D measurements, and compared with 2D measurements, SonoAVC-generated d(V) values tended to be smaller, while SonoAVC-generated m-d values tended to be greater [7]. From the aspect of measurement error, absolute difference of <1 mm was a rather small difference in comparison with the general error of an ultrasound measurement, which was reported to be 1.2 to 3mm in 2D ultrasound [14,15]. Just as Ata et al. stated in their study, it was unlikely that the <1 mm absolute difference in the size of the follicle would affect treatment outcomes in clinic [7]. Nevertheless, our study results gave us a hint that if we used 2D ultrasound and SonoAVC interchangeably, automated SonoAVC might underestimate follicle size.

Furthermore, we found the limits of agreements between 2D diameters and SonoAVC-generated d(V) or m-d values both became wide for large follicle size, that is to say, measurement errors of small follicles were less than that of large follicles. Indeed, it had been shown that larger follicles were associated with greater measurement error, a phenomenon existing in 2D ultrasound examination no matter by an individual observer or by several observers [15,16]. We speculated that for large follicles, measurement error within 2D ultrasound increased even by a single observer in the present study, but measurement error within SonoAVC remained stable, because SonoAVC was theoretically based on the voxel count within the

identified follicle. As a result, the measurement discrepancy between two techniques enlarged for large follicles.

In summary, our study suggested that compared with manual 2D ultrasound, automated SonoAVC might underestimate follicle size, although with small absolute difference <1 mm. The absolute size of a follicle affected the consistency of two techniques.

References

- Raine-Fenning N, Jayaprakasan K, Clewes J (2007) Automated follicle tracking facilitates standardization and may improve work flow. *Ultrasound Obstet Gynecol* 30: 1015-1018. doi:10.1002/uog.5222. PubMed: 18044823.
- 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. PubMed: 18484679.
- Raine-Fenning N, Jayaprakasan K, Chamberlain S, Devlin L, Priddle H et al. (2009) Automated measurements of follicle diameter: a chance to standardize? *Fertil Steril* 91: 1469-1472. doi:10.1016/j.fertnstert.2008.07.1719. PubMed: 18930205.
- Lamazou F, Arbo E, Salama S, Grynberg M, Frydman R et al. (2010) Reliability of automated volumetric measurement of multiple growing follicles in controlled ovarian hyperstimulation. *Fertil Steril* 94: 2172-2176. doi:10.1016/j.fertnstert.2010.01.062. PubMed: 20226448.
- Salama S, Arbo E, Lamazou F, Levailant JM, Frydman R et al. (2010) Reproducibility and reliability of automated volumetric measurement of single preovulatory follicles using SonoAVC. *Fertil Steril* 93: 2069-2073. doi:10.1016/j.fertnstert.2008.12.115. PubMed: 19342038.
- Deutch TD, Joergner I, Matson DO, Oehninger S, Bocca S et al. (2009) Automated assessment of ovarian follicles using a novel three-dimensional ultrasound software. *Fertil Steril* 92: 1562-1568. doi:10.1016/j.fertnstert.2008.08.102. PubMed: 18937942.
- Ata B, Seyhan A, Reinblatt SL, Shalom-Paz E, Krishnamurthy S et al. (2011) Comparison of automated and manual follicle monitoring in an unrestricted population of 100 women undergoing controlled ovarian stimulation for IVF. *Hum Reprod* 26: 127-133. doi:10.1093/humrep/deq320. PubMed: 21112951.
- Raine-Fenning N, Deb S, Jayaprakasan K, Clewes J, Hopkisson J et al. (2010) Timing of oocyte maturation and egg collection: during controlled ovarian stimulation: a randomized controlled trial evaluating manual and automated measurements of follicle diameter. *Fertil Steril* 94: 184-188. doi:10.1016/j.fertnstert.2010.07.719. PubMed: 19342014.
- Raine-Fenning N, Jayaprakasan K, Deb S, Clewes J, Joergner I et al. (2009) Automated follicle tracking improves measurement reliability in patients undergoing ovarian stimulation. *Reprod Biomed Online* 18: 658-663. doi:10.1016/S1472-6483(10)60010-7. PubMed: 19549444.
- Shrout PE (1998) Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 7: 301-317. doi:10.1191/096228098672090967. PubMed: 9803527.
- Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1: 307-310. PubMed: 2868172.
- Myles PS, Cui J (2007) Using the Bland-Altman method to measure agreement with repeated measures. *Br J Anaesth* 99: 309-311. doi:10.1093/bja/aem214. PubMed: 17702826.
- Yavas Y, Selub MR (2009) Ovarian follicular volume and follicular surface area are better indicators of follicular growth and maturation, respectively, than is follicular diameter. *Fertil Steril* 91: 1299-1302. doi:10.1016/j.fertnstert.2008.01.050. PubMed: 18353316.
- Rodríguez-Fuentes A, Hernández J, García-Guzmán R, China E, Iaconianni L et al. (2010) Prospective evaluation of automated follicle monitoring in 58 in vitro fertilization cycles: follicular volume as a new indicator of oocyte maturity. *Fertil Steril* 93: 616-620. doi:10.1016/j.fertnstert.2009.02.058. PubMed: 19376513.
- Forman RG, Robinson J, Yudkin P, Egan D, Reynolds K et al. (1991) What is the true follicular diameter: an assessment of the reproducibility of transvaginal ultrasound monitoring in stimulated cycles. *Fertil Steril* 56: 989-992. PubMed: 1936337.
- Duijkers IJ, Louwé LA, Braat DD, Klipping C (2004) One, two or three: how many directions are useful in transvaginal ultrasound measurement of ovarian follicles? *Eur J Obstet Gynecol Reprod Biol* 117: 60-63. doi:10.1016/j.ejogrb.2004.01.005. PubMed: 15474246.

Author Contributions

Conceived and designed the experiments: PP DY. Performed the experiments: PP XC YL QZ DY. Analyzed the data: PP XC XZ. Contributed reagents/materials/analysis tools: PP XC YL QZ XZ MMABD DY. Wrote the manuscript: PP.