Reproducibility of transvaginal three-dimensional endometrial volume measurements during ovarian stimulation

C.Yaman1,2, M.Sommergruber1, T.Ebner1, W.Pölz2, M.Moser1 and G.Tews1

1Women’s General Hospital, Department of Assisted Reproduction, Lederergasse 47, A-4020 Linz; 2University of Linz, Department for Statistics, Altenbergerstr: 69, A-4040 Linz, Austria

Introduction

Most ultrasound studies published about changes of endometrium during ovarian stimulation have measured the endometrial thickness in two-dimensional planes. There is no consensus on the role of endometrial thickness for prediction of pregnancy rates in assisted reproduction. Several investigators have suggested that endometrial thickness is a useful criterion in predicting embryo implantation and pregnancy rate (Gonen and Caspar, 1990; Noyes et al., 1995; Remohi et al., 1997). However, the relationship between the endometrial thickness diagnosed by two-dimensional ultrasound and pregnancy rate is not universally accepted (Fleischer et al., 1986; Imedemhe et al., 1987; Strohmer et al., 1994; Oliveira et al., 1997; Ohno and Fujimoto, 1998). In addition, the calculation of endometrium volume using two planes does not play an important role for routine use in infertility treatment.

Before evaluating the clinical role of a method, its reproducibility has to be checked. Only a few studies describing the clinical application of transvaginal 3D-ultrasound volume measurements are available (Gregg et al., 1993; Steiner et al., 1993; Brunner et al., 1995; Mensah et al., 1996; Riccabona et al., 1996; Lee et al., 1997). Only one of these studies reported about endometrial volume determination based on transvaginal 3D-ultrasound in stimulated cycles (Mensah et al., 1996).

The aim of this study was to document the reproducibility of endometrial volume measurement by 3D-ultrasound.

Materials and methods

One three-dimensional (3D) endometrial measurement was performed on each of 57 consecutive infertile patients undergoing ovarian stimulation for in-vitro fertilization (IVF) and embryo transfer or intracytoplasmic sperm injection (ICSI). In all cases ovarian stimulation was carried out according to the long protocol (Smitz et al., 1987) using a gonadotrophin-releasing hormone agonist in association with human menopausal gonadotrophin (HMG). The volume measurements were done between day 5 and day 12 (mean ± SD = 8.24 ± 1.68) of ovarian stimulation.

All 3D scans were obtained using the Voluson 530 D (Kretztechnik AG, Zifp, Austria) with a transvaginal 5–7.5MHz volume transducer. After ultrasonographic investigation of the uterus morphology, the endometrium was visualized in B mode. Plane B was chosen in preference to A or C after preliminary monitoring. No difference was expected between volume measurements in different planes. Then the system was switched into ‘volume mode’. After targeting of the endometrium in a vertical plane by the volume box, the slow volume acquisition setting was activated. Volume data were captured by holding the transducer stationary while its crystal electronically swept up to 95° for 2–3 s. All scanned volumes were stored on a 540 Mbyte hard disc with an integrated magneto-optical drive for later measurements and analysis. Two sets of measurements were made by observer II to calculate intra-observer reliability.

Two different methods were used for volume measurements; the full planar (contour) and the three distance method in which the volume was computed from several parallel sections. Measurements with the three distance method were performed in fundal height, where the endometrium expanded (Figure 1). In the full planar method the volume was computed from several parallel sections of the endometrium to the level of internal os (Figure 2). Mean number of parallel sections for the full planar method was 8.426 and for the full planar method 0.9565. Intra-observer reliability for the three distance method was 0.8426 and for the full planar method 0.9565. Interobserver reliability for the three distance method was 0.6667 and for the full planar method 0.9394. The correspondence between and within observers seemed to be good. Both methods are reliable, but the full planar method seems to provide slightly better reproducibility in regard to endometrium volume measurement.

Key words: endometrium volume/IVF/reproducibility/three-dimensional ultrasound
Figure 1. Three distance method of endometrial volume measurement in a normal uterus. Top left, transversal plane (A-plane); top right, vertical plane (B-plane); bottom left, frontal plane or depth of endometrium (C-plane).

Figure 2. Full planar method of endometrial volume measurement demonstrated in a case of uterus arcuatus: A-plane, tracing of endometrium in a transversal plane; B-plane, tracing of endometrium for planimetric (full planar method) measurements; C-plane, arrow showing the convex nature of fundal endometrium.

(Kolmogorov–Smirnov test with Lilliefors corrections). As no significant deviation from normality was found, the method of intraclass correlation was used to express inter- and intra-observer-reliability (Armitage and Berry, 1994).

The intraclass correlation coefficient is defined as the correlation between any two measurements on the same object. Its value lies between 0 and 1, where 1 indicates total reliability. In this study it is a measure of correlation between the endometrial volume measurements obtained by any two randomly chosen observers on the same individual. A high correlation between the measurements...
of two observers alone is not enough to prove interobserver reliability. The regression line must also pass through the origin and the slope should be close to 1.

Ultrasound examination was performed routinely, therefore the study was not subject to the institutional review board approval. The examinations and the measurements were done by two experienced observers in pelvic ultrasonography. They analysed the stored data independently and were blinded to their own and to the other observer’s results.

**Results**

The mean endometrium volumes measured from the first observer with the full planar method and with the three distance method were 4.25 and 4.90 ml respectively. The corresponding measurements from the second observer were 4.07 and 3.79 ml, and 4.16 and 3.59 ml. The mean age of the patients was 34.07 ± 5.63 years. The mean number of ampoules HMG used was 27.70 ± 12.43. The mean number of ampoules HMG per day was 3.43 ± 1.56. The mean duration of stimulation was 8.24 ± 1.28 days.

**Whole group (n = 57)**

Statistical data for the calculation of intraclass correlation coefficients are given in Table I. These data yielded an interobserver-reliability index for the three distance method of 0.6667 and for the full planar method of 0.9565, whereas the intra-observer reliability index for the three distance method was 0.8426 and for the full planar method 0.9394 (Table I).

A paired t-test showed there was no statistical difference between the two methods. A linear regression using the full planar method as the independent and the three distance method as the independent variable, yielded the following results (Figure 3): intercept = 0.348 (not statistically different from 0), slope = 0.962 (statistically different from 1, P < 0.01). When data were plotted as boxplots (Figure 4) or scatterplots (Figure 5), it was clear that volume measurements obtained by the two methods were very similar.

**Cases with abnormal uterus or poor endo-myometrial contrast (n = 7)**

In seven cases out of the total collective group (Figure 2), the interobserver reliability for the three distance method was only 0.2098 and for the full planar method 0.9418. Intra-observer reliability for the three distance method was 0.7869 and for the full planar method 0.9602.

**Discussion**

Two-dimensional ultrasound is a reliable method for determining size and morphology of the structures in the female pelvis. The examiner obtains two-dimensional sectional images, from which morphology, size, pathological changes and volume can be identified. A mean absolute error rate of 12.6% for two-
Endometrium volume measurements by 3D-ultrasound

Figure 5. Scatterplots. First observer (\(i^1\)), second observer (\(i^2\)). (a) Full planar method (intra-observer ratings): \(y = 0.279 + 0.911x\). (b) Three distance method (intra-observer ratings): \(y = 0.547 + 0.904x\). (c) Comparison of the three distance method with full planar method (first observer): \(y = 0.301 + 0.806x\). (d) Comparison of the three distance method with full planar method (second observer): \(y = 1.244 + 0.746x\).

Three-dimensional imaging is not a new technique. Three-dimensional imaging by computerized tomography and magnetic resonance imaging have led to significant improvements in medical diagnosis and treatment in recent years. It provided a stepping stone to the development of safe and inexpensive 3D ultrasound. This method also allows spontaneous measurement of a third plane (frontal plane = the depth or right–left diameter) of endometrium. Accordingly, endometrial growth can be monitored with more accuracy. The absolute error rate for 3D volume measurement (Riccabona et al., 1996) was only 6.4%.

Two methods are available for transvaginal 3D-ultrasound volume measurements. Our results showed that both the full planar and three distance methods had high reproducibility. Agreement between and within the observers was good; the mean difference and the variation in the difference were small compared with the total mean with standard deviation (Table I). Therefore it was considered acceptable to combine the assessments of the two observers and to use the mean value in the main analysis. This was the most important finding of this study. The reproducibility of volume measurements in a normally configured uterus with physiological echo-contrast of endometrium to myometrium was higher than that in a subgroup of patients (\(n = 7\)) with asymmetrically shaped endometrium, either uterus arcuatus (\(n = 2\)), hyperflexion of uterus (\(n = 1\)), left-deviated uterus (\(n = 1\)), or uterus with unclear endomyometrial echo-contrast (\(n = 3\)).

Neither the full planar nor the three distance method can give exact calculations of volume changes with correlation coefficient 1. The reproducibility (interobserver reliability) of the full planar method measurements was slightly better but the results were statistically not different from the three

dimensional volume measurement has been described (Riccabona et al., 1996). It is probable that the two-dimensional endometrium volume measurement is not widely practised because of its limited accuracy.

...
distance method \((P < 0.116)\). In a subgroup of seven patients with asymmetrically shaped endometrium and endometrium with poor contrast to myometrium, the two methods revealed significantly different volumes \((P < 0.012)\). In this subgroup, extremely low agreement was diagnosed between the two observers (correlation coefficient 0.2098).

Generally, the endometrium showed a good contrast to myometrium, so measurements could be performed easily. In cases with poor contrast, it was difficult to identify the internal os for measurements of the length of endometrium in the vertical plane.

An important criterion for the three distance method is the height in which volume measurements will be performed. Volumes captured from the fundal part of the endometrial cavity were always greater than those from the level of the internal os. Therefore, to evaluate the reproducibility of the three distance method, measurements should be performed in the same height of endometrium. The reason for the better reproducibility of the full planar method was that several parallel plane sections could be measured consecutively. This minimized the rate of error that occurred especially in volume measurements of asymmetrically shaped uteri. However, measurements with the full planar method took three to four times as long, which constituted the biggest disadvantage of this method.

The stored volume data can be analysed days, weeks, or years later. Effective use of this system includes scanning experience. The investigator must have sufficient experience to capture satisfactory volume data for storage. The possibility of retrospective verification is also of forensic interest. Furthermore, it is not confined to the investigator, but allows additional consultation by other specialists. In this way, more and better information can be obtained and this information can be useful in optimizing patient management.

Uterine receptivity to the embryo is vital for successful implantation and pregnancy. Whether ultrasound evaluation of the endometrium can distinguish between unfavourable and favourable endometrium has yet to be determined. No correlation has been found between pre-ovulatory endometrial thickness, echo pattern, and fecundity (Ohno and Fujimoto, 1998). Neither were any relationships found between endometrial appearance, endometrial steroid receptors, and steroid hormone concentrations in serum. However, a positive correlation between endometrial thickness and implantation has been reported (Remohi et al., 1997), as well as between endometrial thickness and serum estradiol. Based on the strong correlation between uterine dimension and endometrial thickness, it was suggested that endometrial thickness is determined by individual uterine architecture and is therefore not predictive of the likelihood of implantation (Strohmer et al., 1994). A good correlation between intra-observer and interobserver measurements of endometrial thickness has been found (Spandorfer et al., 1998). Nevertheless, the ability to quantify the volume of the endometrium using 3D ultrasound may improve the prediction of pregnancy in assisted reproduction treatment, because the outcome can be related to endometrial volume, which may provide more precise data than two-dimensional measurement of endometrial thickness. In addition, preliminary data (Gruboeck et al., 1996) suggested that the measurement of endometrium volume was superior to that of endometrial thickness as a diagnostic test for the detection of endometrial cancer.

In conclusion, our data demonstrate that volume estimation of the endometrium during ovarian stimulation can be performed by transvaginal 3D ultrasound. Both full planar and three distance methods had a high degree of reproducibility. Overall, the reproducibility of the full planar method was slightly better. Therefore the full planar method may be recommended for clinical studies of endometrium volume estimation. Whether the volume estimation by transvaginal 3D ultrasound can be used in routine management of patients undergoing IVF or ICSI has to be the subject of further clinical studies.

References


Received on April 20, 1999; accepted on July 12, 1999