Human Reproduction Vol.17, No.12 pp. 3110–3113, 2002

Endometrial volume and thickness measurements predict pituitary suppression and non-suppression during IVF

Tim J.Child[1], Camille Sylvestre and Seang Lin Tan

McGill Reproductive Center, Department of Obstetrics and Gynecology, Royal Victoria Hospital, McGill University, Montreal, Quebec, H3A 1A1, Canada

[1]To whom correspondence should be addressed at: Oxford Fertility Unit, Level 4, Women's Centre, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK. E-mail: Timothychild@yahoo.com

BACKGROUND: The aim of this prospective study was to evaluate the usefulness of three-dimensional (3D) ultrasound measurement of endometrial volume and thickness as a predictor of pituitary suppression and non-suppression following GnRH agonist administration for IVF. METHODS: A total of 144 women undergoing 164 IVF cycles had transvaginal ultrasound measurement of their endometrial volume and thickness following 8–14 days of buserelin acetate administration. Serum estradiol concentrations were measured on the same day. Receiver operating characteristic (ROC) curve analysis was used for statistics. A ROC curve was produced for each of four estradiol thresholds commonly used by clinics to diagnose pituitary suppression (100, 150, 200, 250 pmol/l). From each curve, endometrial volume and thickness thresholds that best predicted pituitary suppression and, separately, non-suppression were selected and the associated sensitivity, specificity, positive and negative predictive values were reported. RESULTS: The area under the curve (AUC) was consistently higher (better test) for 3D volume than thickness estimation for all four estradiol thresholds, although it was only significantly different when a threshold of 200 pmol/l was used. The AUC increased towards 1.0 (perfect test) for both volume and thickness measurement as the selected estradiol threshold increased. Very different volume and thickness thresholds were optimal depending on whether the aim of the test was to predict pituitary suppression or non-suppression. CONCLUSIONS: 3D endometrial volume estimation provides a new tool, alongside endometrial thickness measurement, to diagnose pituitary suppression and non-suppression during IVF. Different endometrial thresholds must be selected depending upon whether the priority is to identify pituitary suppressed, or arguably more importantly, non-suppressed cycles.

Key words: endometrial thickness/endometrial volume/pituitary suppression/serum estradiol/3D-ultrasound

Introduction

IVF and embryo transfer is an established and successful form of treatment for infertility. High pregnancy rates are achieved using ovarian stimulation, since this increases the numbers of oocytes collected and consequently embryos available for transfer (Templeton and Morris, 1998). Data from randomized controlled trials demonstrate superior IVF outcomes when ovarian stimulation is commenced following pituitary suppression (Tan *et al.*, 1994). Pituitary suppression is generally achieved by administering GnRH agonists as part of a 'long' protocol. Pituitary suppression results in a hypo-estrogenic state that may be confirmed through measurement of the serum estradiol concentration (Ibrahim *et al.*, 1990). However, no consensus exists on the optimum degree of hypo-estrogenism to achieve prior to starting ovarian stimulation. Common thresholds for pituitary suppression used by clinics include 100, 150 or 200 pmol/l.

The endometrium acts as a bioassay for circulating serum estradiol (Nakamura *et al.*, 1996). Ultrasonographic measurement of endometrial thickness may therefore be used as a marker for serum hyper- or hypo-estrogenism and can be used to confirm pituitary suppression prior to ovarian stimulation during IVF treatment (Barash *et al.*, 1998). The advantages of this approach over routine serum estradiol testing in all patients are that (i) the number of blood tests required during an IVF cycle is reduced and (ii) ultrasound imaging identifies pelvic pathology, including ovarian cysts or endometrial polyps, that may adversely affect cycle outcome. Surprisingly, though widely used as a predictor of pituitary suppression, only two studies have examined the value of ultrasound measurement of endometrial thickness or volume in screening for pituitary down-regulation during IVF (Barash *et al.*, 1998; Yaman *et al.*, 2000). Both studies focused on selecting endometrial thresholds that best identified women who were pituitary suppressed, not those with failed suppression. It is arguably more important

to identify pituitary non-suppressed patients, since they may be at increased risk of a blunted follicular response to ovarian stimulation.

Yaman and colleagues examined the role of three-dimensional (3D) endometrial volume estimation in predicting pituitary suppression in an IVF programme (Yaman *et al.*, 2000). However, their study included 46 patients of whom only four had serum estradiol concentrations >60 pg/ml, and positive and negative values of the test were not calculated. In particular, the aim of their study was to identify pituitary suppressed patients rather than women with failed suppression.

The aim of the present study is to extend the observations of Yaman *et al*. with a much larger group of 164 cycles in order to document the role of 3D endometrial volume and thickness estimation in predicting both failed and successful pituitary suppression.

Materials and methods

The study group comprised 144 women undergoing 164 IVF cycles. The median (range) age of patients was 34 years (24–44). The primary infertility diagnosis was unexplained (32%), male factor (31%), tubal damage (26%), endometriosis (8%) and other causes (3%). Women underwent a long GnRH agonist protocol with an early follicular phase start. Patients were pretreated with either the oral contraceptive pill (Marvelon; Organon, Scarborough, Ontario, Canada) or oral progesterone (Norlutate; Parke-Davis, Scarborough, Ontario, Canada). Pill-pretreated women took Marvelon for 2 weeks from day 1 of menstruation (Biljan *et al.*, 1998). The pill was then stopped and GnRH agonist treatment commenced. Progesterone-pretreated women took Norlutate for 5 days from day 1 of menstruation (Engmann *et al.*, 1999a). GnRH agonist was added from day 2 of menstruation. Pituitary suppression was assessed after 14 days of GnRH agonist treatment in the Marvelon group and after 8–14 days in the Norlutate group. The GnRH agonist used in all cases was a once daily s.c. injection of buserelin acetate 500 μg.

Ultrasound scans were performed using a machine with two-dimensional (2D) and 3D capabilities (Voluson 530D, Medison Inc., CA, USA) attached to a 5–7.5 MHz transvaginal probe. Scans were performed by one of two experienced ultrasonographers: T.J.C. or C.S. Patients were asked to empty their bladder prior to examination. First, a 2D ultrasound scan was performed in order to measure endometrial thickness. The uterus was visualized in the longitudinal plane, ensuring that the complete endometrial echo from cervix to fundus was clearly seen. The endometrial thickness was taken as the maximum distance between the two myometrial–endometrial interfaces at right angles to the cavity. The machine was then switched to volume mode and the slow sweep speed selected for maximum image resolution. The volume box was placed over the endometrium. Volume acquisition was commenced and the transducer held still for a few seconds during the crystal sweep. The scanned volumes were stored on the machine's hard drive for later analysis.

The contour method was used to measure the endometrial volume. We have previously demonstrated high reliability and reproducibility of 3D endometrial volume measurements using this method (Kyei-Mensah *et al.*, 1996), as have subsequent investigators (Yaman *et al.*, 1999). All measurements were performed by one investigator (T.J.C.). The volume was computed by the ultrasound machine from several parallel cross-sections from the fundal part of the endometrium to the internal os. The intra-observer reliability of endometrial thickness and volume estimation was calculated by measuring the endometrial thickness and volume of 10 patients three times. The intra-class correlation coefficients were calculated as 0.96 (volume) and 0.92 (thickness). These results indicate high intra-observer reliability.

Following the ultrasound scan, all patients had a blood sample taken for estradiol assay. This was performed using the ACS:180 Estradiol-6 II assay (Bayer Corporation, Tarrytown, NY, USA). The assay had between-runs coefficients of variation (CV) of 27.6% at a mean of 77 pmol/l, 15.3% at 191 pmol/l, and 7.5% at 352 pmol/l estradiol concentrations. The endometrial volume for each patient was calculated before running the estradiol assay.

Statistical analysis was performed using receiver operating characteristic (ROC) curve analysis. The ROC curve represents the probability of true positive results (sensitivity) as a function of the probability of false positive results (1–specificity). In order to calculate each curve, a particular threshold value for the state variable (serum estradiol concentration) must be selected. This is the arbitrary value at which pituitary suppression is defined. Since no consensus exists on the optimal degree of pituitary suppression as measured by serum estradiol concentration, we calculated ROC curves for each of four commonly used thresholds; 100 pmol/l, 150 pmol/l, 200 pmol/l and 250 pmol/l. For each estradiol threshold, separate curves were produced for volume and thickness measurements, making a total of eight ROC curves. The sensitivity and specificity of a particular endometrial volume (or thickness) cut-off, as a screening test for pituitary suppression or non-suppression, changes depending on the selected estrogen threshold. The possible combinations of sensitivity and specificity obtained when varying the endometrial cut-off point for a particular estradiol threshold are combined into an area under the curve (AUC). The AUC measures how good (AUC close to 1.0) or poor (AUC close to 0.5) a test is. To compare 2D and 3D endometrial measurement as a test, the AUC for thickness is compared against the AUC obtained for volume measurements at the same estradiol threshold. Four such comparisons were performed using the statistical method of Hanley and McNeil (Hanley and McNeil, 1983), i.e. volume versus thickness for each of the concentrations 100, 150, 200, and 250 pmol/l estradiol. Analyses were performed using SPSS (SPSS, Chicago, IL, USA).

Results

A total of 164 cycles was included in the study. In all, 41 (25%) women had serum estradiol concentrations >100 pmol/l, 31 (19%) >150 pmol/l, 21 (13%) >200 pmol/l, and 16 (10%) had estradiol concentrations >250 pmol/l on the day of examination. Eight ROC curves were produced, one for volume and one for thickness, for each of the four estradiol thresholds. The AUC of each curve is listed in the third column of Table I. Table I illustrates the test characteristics for volume and thickness thresholds chosen by the authors for each estradiol curve to best diagnose pituitary suppression. Any endometrial threshold (column 4) could have been chosen. As the chosen endometrial threshold increases, the proportion of pituitary suppressed cycles identified (sensitivity) will increase but the proportion of non-suppressed cycles (specificity) will decrease.

Table II illustrates those measurements selected by the authors to best diagnose failure of pituitary suppression. Note that the AUCs for each estradiol threshold and endometrial measurement method are the same as in Table I, since the ROC curves are the same. The authors have merely used the same

Table I. Use of endometrial volume and thickness measurements to diagnose pituitary suppression at four commonly used estradiol thresholds

Estradiol threshold (pmol/l)	Endometrial Parameter	AUC	Measurement	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
100	Volume	0.72	<1.48 ml	91.9	34.1	80.7	58.3
	Thickness	0.69	<5.9 mm	91.9	34.1	80.7	58.3
150	Volume	0.80	<1.48 ml	91.7	41.9	87.1	54.2
	Thickness	0.75	<5.9 mm	91.7	41.9	87.1	54.2
200	Volume	0.88	<1.48 ml	91.5	54.5	92.9	54.2
	Thickness	0.80	<6.0 mm	93.7	54.5	93.0	57.1
250	Volume	0.92	<1.48 ml	91.8	70.6	96.4	54.2
	Thickness	0.91	<6.0 mm	93.9	64.7	95.8	52.4

AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value.

Table II. Use of endometrial volume and thickness measurements to diagnose failure of pituitary suppression at four commonly used estradiol thresholds

Estradiol threshold (pmol/l)	Endometrial parameter	AUC	Measurement	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
100	Volume	0.72	≥0.50 ml	80.5	45.5	33.0	87.5
	Thickness	0.69	≥3.3 mm	75.6	44.7	31.3	84.6
150	Volume	0.80	≥0.50 ml	90.3	45.9	28.0	95.3
	Thickness	0.75	≥3.7 mm	80.6	54.5	29.1	92.3
200	Volume	0.88	≥0.83 ml	90.9	73.2	34.5	98.1
	Thickness	0.80	≥4.0 mm	81.8	63.4	25.7	95.8
250	Volume	0.92	≥1.02 ml	94.1	80.3	34.8	99.2
	Thickness	0.91	≥4.5 mm	94.1	77.6	32.7	99.1

AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value

curve but selected endometrial thresholds that have an optimal sensitivity and specificity for predicting failure of pituitary suppression.

The AUC for volume or thickness measurement increases as the selected estradiol threshold chosen to denote pituitary suppression increases (Table I). The AUC of the volume curve is also consistently higher than that of endometrial thickness (column 3). However, the difference in AUC between volume and thickness measurement only reached statistical significance when using an estradiol threshold of 200 pmol/l ($P < 0.05$).

Discussion

It is generally accepted that pituitary suppression should be achieved prior to starting gonadotrophin ovarian stimulation, though the optimal degree of suppression remains to be determined. Although women may have suppression confirmed through serum estradiol assay, there are advantages in ultrasound confirmation, including the ability to diagnose ovarian or endometrial pathology. In addition, ultrasound measurement on the day of pituitary suppression of the ovarian volume and antral follicle count (T.J.Child, C.Sylvestre and S.L.Tan, unpublished data), and ovarian stromal blood flow (Engmann *et al.*, 1999b) are predictive of the follicular response and may be used to select the most appropriate starting dose of gonadotrophin.

Previous studies have attempted to determine the endometrial thickness or volume thresholds that best predict pituitary suppression rather than non-suppression. The difference is subtle but important. For instance, from the data of Yaman and colleagues (Yaman *et al.*, 2000), an endometrial volume of 1.8 ml has a sensitivity and specificity for diagnosing pituitary suppression (when defined as serum estradiol <40 pg/ml, equivalent to ~150 pmol/l) of 91.2 and 33.3% respectively. The use of this volume threshold means that only 33.3% of non-suppressed women will be identified. Put another way, two-thirds of pituitary non-suppressed women would fall into the 'low' endometrial volume range on ultrasound and ovarian stimulation would be commenced unnecessarily.

Our data suggest that different endometrial thickness or volume thresholds should be selected, depending on whether one wishes to screen for suppression or non-suppression. For example, if suppression is considered as an estradiol concentration <150 pmol/l, then an endometrial thickness of <5.9 mm will identify (sensitivity) 91.7% of suppressed patients (Table I). However, only 41.9% (Table I) of non-suppressed women will have an endometrial thickness of >5.9 mm and be identified as having failed pituitary suppression. For the same estradiol threshold of 150 pmol/l, an endometrial thickness of ≥3.7 mm will identify 80.6% of non-suppressed patients (Table II). If a higher sensitivity is preferred (higher likelihood of identifying non-suppressed women), then the endometrial threshold thickness may be reduced further. However, the positive predictive value will also reduce, meaning that more patients with an endometrium thicker than the threshold are in fact suppressed and will undergo a needless blood test.

When selecting endometrial thresholds for screening for pituitary suppression, there appears little advantage of endometrial volume over thickness measurement (Table I), since the sensitivities and specificities between the two methods are similar. This confirms the findings of Yaman and colleagues (Yaman *et al.*, 2000).

When screening for failed suppression, 3D volume estimation achieved higher test sensitivities for similar test specificities (Table II). This could be because in comparison with pituitary suppressed patients, those with failed suppression have a greater endometrial mass. Measurement of high endometrial mass could conceivably be more accurately estimated by volume rather than thickness measurement. However, endometrial volume measurement was only shown to be significantly superior to thickness measurement (AUCs significantly different) when using an estradiol threshold of 200 pmol/l.

Though endometrial volume measurement is simple to perform, the calculation takes longer than the measurement of the endometrial thickness. The volume is calculated from serial tracings of the endometrial outline from internal cervical os to fundus. In addition, when the endometrium is very thin, it may be difficult to identify the myometrial–endometrial interface when tracing the endometrial outline on the serial cross-sections. Each measurement takes between 3 and 5 min.

In summary, the current data support and extend previous work suggesting that 3D endometrial volume estimation provides a new tool in assessing the degree of pituitary suppression during IVF treatment. When selecting endometrial thresholds to screen for pituitary suppression, the sensitivities and specificities obtained were similar between volume and thickness measurements. However, when selecting endometrial thresholds to screen for failure of pituitary suppression, which is perhaps of greater clinical use, 3D volume estimation performed slightly better as a test. Ultrasonographic measurement of endometrial volume or thickness can replace routine measurement of serum estradiol as a predictor of the state of pituitary suppression.

References

Barash, A., Weissman, A., Manor, M., Milman, D., Ben-Arie, A. and Shoham, Z. (1998) Prospective evaluation of endometrial thickness as a predictor of pituitary down-regulation after gonadotropin-releasing hormone analogue administration in an in vitro fertilization program. *Fertil. Steril.*, **69**, 496–499.

Biljan, M.M., Mahutte, N.G., Dean, N., Hemmings,. R, Bissonnette, F. and Tan, S.L. (1998) Effects of pretreatment with an oral contraceptive on the time required to achieve pituitary suppression with gonadotropin-releasing hormone analogues and on subsequent implantation and pregnancy rates. *Fertil. Steril.*, **70**, 1063–1069.

Engmann, L., Maconochie, N., Bekir, J. and Tan, S.L. (1999a) Progestogen therapy during pituitary desensitization with gonadotropin-releasing hormone agonist prevents functional ovarian cyst formation: a prospective, randomized study. *Am. J. Obstet. Gynecol.*, **181**, 576–582.

Engmann, L., Sladkevicius, P., Agrawal, R., Bekir, J.S., Campbell, S. and Tan, S.L. (1999b) Value of ovarian stromal blood flow velocity measurement after pituitary suppression in the prediction of ovarian responsiveness and outcome of in vitro fertilization treatment. *Fertil. Steril.*, **71**, 22–29.

Hanley, J.A. and McNeil, B.J. (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, **148**, 839–843.

Ibrahim, Z.H., Matson, P.L., Buck, P., Critchlow, J.D., Newman, M.C., Horne, G., Hughes, S. and Lieberman, B.A. (1990) Use of buserelin in an IVF programme for pituitary-ovarian suppression prior to ovarian stimulation with exogenous gonadotrophins. *Hum. Reprod.*, **5**, 258–262.

Kyei-Mensah, A., Maconochie, N., Zaidi, J., Pittrof, R., Campbell, S. and Tan ,S.L. (1996) Transvaginal three-dimensional ultrasound: reproducibility of ovarian and endometrial volume measurements. *Fertil. Steril.*, **66**, 718–722.

Nakamura, S., Douchi,T., Oki, T., Ijuin, H., Yamamoto, S. and Nagata, Y. (1996) Relationship between sonographic endometrial thickness and progestin-induced withdrawal bleeding. *Obstet. Gynecol*., **87**, 722–725.

Tan, S.L., Maconochie, N., Doyle, P. Campbell, S., Balen, A., Bekir, J., Brinsden, P., Edwards, R.G. and Jacobs, H.S. (1994) Cumulative conception and live-birth rates after in vitro fertilization with and without the use of long, short, and ultrashort regimens of the gonadotropin-releasing hormone agonist buserelin. *Am. J. Obstet. Gynecol.*, **171**, 513–520.

Templeton, A., and Morris, J.K. (1998) Reducing the risk of multiple births by transfer of two embryos after in vitro fertilization. *N. Engl. J. Med.*, **339**, 573–577.

Yaman, C., Sommergruber, M., Ebner, T., Polz, W., Moser, M. and Tews, G. (1999) Reproducibility of transvaginal three-dimensional endometrial volume measurements during ovarian stimulation. *Hum. Reprod.*, **14**, 2604–2608.

Yaman, C., Ebner, T., Sommergruber, M., Hartl, J., Polz, W. and Tews, G. (2000) Three-dimensional endometrial volume estimation as a predictor of pituitary down-regulation in an IVF-embryo transfer programme. *Hum. Reprod*., **15**, 1698–1702.

Submitted on April 15, 2002; accepted on August 19, 2002