Comparison of automated and manual follicle monitoring in an unrestricted population of 100 women undergoing controlled ovarian stimulation for IVF

Baris Ata*, Ayse Seyhan, Shauna Leigh Reinblatt, Einat Shalom-Paz, Srinivasan Krishnamurthy, and Seang Lin Tan

McGill Reproductive Centre, Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, McGill University, Montreal, QC, Canada

*Correspondence address. Royal Victoria Hospital, 687 Pine Avenue West, Room F6.58, Montreal, QC, Canada H3A 1A1. E-mail: baris.ata@muhc.mcgill.ca

Submitted on August 17, 2010; resubmitted on October 14, 2010; accepted on October 19, 2010

BACKGROUND: Ovarian response to gonadotrophin stimulation is monitored with serial ultrasound (US) examinations. Sonography-based Automated Volume Count (SonoAVC) is a relatively new three-dimensional (3D) US technology, which automatically generates a set of measurements including the mean follicular diameter (MFD) and a volume-based diameter ($d_v$) for each follicle in the ovaries. The present study aimed to assess the applicability and reproducibility of this automated follicle measurement method in an IVF programme.

METHODS: For this prospective method comparison study, 100 women undergoing US monitoring of a controlled ovarian stimulation cycle were recruited. Each follicle was manually measured by taking the mean of maximal diameters on three orthogonal planes with two-dimensional (2D) US. A 3D volume of each ovary was then captured. The ovarian volumes were later analysed using SonoAVC. The agreement between the two methods for the numbers of follicles and the size of the leading follicle was assessed with the Bland–Altman method. The reproducibility of SonoAVC measurements was assessed with the intraclass correlation coefficient (ICC).

RESULTS: Both SonoAVC-generated MFD and $d_v$-based follicle counts, as well as the leading follicle diameter, had good agreement with conventional 2D US measurements. SonoAVC measurements had very good reproducibility, with ICC $\geq 0.8$ for most evaluations.

CONCLUSIONS: Automated follicle monitoring with SonoAVC can replace or be used interchangeably with conventional 2D measurements. Automated follicle monitoring can save time, provide a method of quality control and create opportunities for developing HCG criteria based on follicular volume or for monitoring patients from a distance.

Key words: SonoAVC / automated volume calculation / automated ultrasound / follicle monitoring / follicle tracking

Introduction

Serial ultrasound (US) examinations are performed to assess the number and size of follicles during controlled ovarian stimulation (COS) for IVF. The likelihood of a follicle containing a mature oocyte is related to its size (Rosen et al., 2008). Accordingly, oocyte collection is done when as many follicles as possible in the growing cohort are large enough to suggest the presence of mature oocytes.

Traditionally, the size of a follicle is assessed by measuring its diameter with two-dimensional (2D) US. However, a follicle is a three-dimensional (3D) structure and its volume is the most accurate measure of its size. Using the diameter as a surrogate for volume assumes the follicles are spheres. However, in the case of multifollicular growth, follicles rarely attain a spherical conformation and most are ellipsoids or have irregular shapes. Therefore, the diameter of a follicle is an imperfect surrogate for its true size (Penzias et al., 1994; Raine-Fenning et al., 2009a,b). Moreover, there is no universal standard for measuring the follicular diameter (Raine-Fenning et al., 2008). Some clinics use the single largest diameter, whereas others calculate the mean of two, three or four diameters, measured in one or two planes as a surrogate of the true follicular size. Identification of these diameters is subjective and contributes to interobserver variability. The reliability of follicular measurements decreases as the number of follicles increases (Forman et al., 1991; Penzias et al., 1994). It has been shown that manual measurement of follicles with 2D US is often inaccurate and subject to significant intra-
interobserver variability (Ritchie, 1985; Forman et al., 1991; Penzias et al., 1994).

Sonography-based Automated Volume Count (SonoAVC; GE Medical Systems, Kretz, Austria) is a new software, which automatically identifies the follicles in a captured ovarian volume, and provides estimates of their dimensions. Previous studies have demonstrated that SonoAVC accurately measures follicular volume (Raine-Fenning et al., 2009a,b; Lamazou et al., 2010; Salama et al., 2010). Studies assessing the agreement between SonoAVC and 2D measurements in women undergoing IVF reported good agreement between the methods (Deutch et al., 2009; Raine-Fenning et al., 2009a,b). However, only a single follicle or a fraction of all available follicles were studied at the discretion of the sonographer, and the measurements were done on a fixed day of the stimulation cycle.

The present study aimed to compare SonoAVC-generated follicular measurements with 2D measurements obtained on various days of a treatment cycle in an unrestricted population of women undergoing COS-SVF. We also assessed the interobserver variation in SonoAVC treatment cycle in an unrestricted population of women undergoing COS-SVF. We also assessed the interobserver variation in SonoAVC follicle measurements in order to evaluate its reliability in clinics where several sonographers scan patients during the course of a single treatment cycle.

**Materials and Methods**

This prospective method evaluation study was approved by the Ethics Committee of the Royal Victoria Hospital and all participants gave informed consent.

One hundred women undergoing COS-IVF at McGill Reproductive Centre between November 2009 and February 2010, who had received gonadotropins for ≥5 days and had been scanned by the same physician (B.A.) using a Voluson E8 Expert (GE Healthcare) with a 5–9 MHz intravaginal probe, were recruited without any exclusion criteria. Each woman contributed to the study with one US examination in order to avoid multiplicity associated with multiple observations on the same subject.

**Conventional 2D scans**

Each ovary was systematically scanned by scrolling from one pole to the other. Each follicle was assessed by measuring the maximal diameters on three orthogonal planes. The mean of three diameters was calculated. As per the study protocol, only the 2D US results were used for clinical management of the COS cycles.

**3D scans and SonoAVC analysis**

Upon completion of 2D measurements, a 3D volume of each ovary was captured without performing any automated measurements. The raw ovarian volumes were labelled with an identification number and stored. One month after data collection, the ovarian volumes were analysed with SonoAVC in a random order to prevent the 2D US results being remembered.

SonoAVC automatically identifies hypoechoic follicles within the captured ovarian volume and generates a set of measurements for each follicle. These measurements include the largest diameters in three orthogonal planes, the mean follicular diameter (MFD), the volume of the follicle and the volume-based diameter ($d_v$) of the follicle. The volume calculation is based on the voxel count within the identified follicle. It therefore represents a true measure of follicular volume. Although the MFD is the arithmetic mean of the three longest orthogonal diameters, $d_v$ is the diameter of a perfect sphere with the same volume as the follicle that is measured. After calculation of the actual volume of a follicle, SonoAVC calculates the diameter of a perfect sphere which has the same volume as the follicle by using the relaxed sphere diameter formula.

Upon completing automated analysis and prior to generating the final report, SonoAVC provides post-processing options. Briefly, ovarian images are manually rotated 360° to identify any errors in automatic identification. Any follicles which could be overlooked by the software can be added to, and any hypoechoic regions such as free fluid around the ovaries or blood vessels adjacent to the ovaries which could be erroneously included in the follicle count can be excluded from the follicle count by simply clicking on them using the ‘add/remove’ function. The ‘cut’ function is used to separate any adjacent follicles with thin follicular walls, which could have been erroneously identified and measured as a single follicle by the software, and to trim follicular borders to fit to the exact shape of the follicle. Rarely, a single follicle with heterogenous echogenicity can be identified as separate follicles by the software, and the ‘merge’ function is used to combine them to be counted as a single follicle. The settings of growth and separation within the software were kept uniform at default values of ‘mid’ for all follicle measurements.

Follicles and their corresponding measurements are colour coded for easy interpretation of the worksheet (Fig. 1).

**Statistical considerations**

The multifollicular cohort represents an unusual type of variable to be analysed. A complete clinical assessment of follicular growth includes the distribution of follicles across different categories based on size and the size of the leading follicle, which is another important piece of information. While the latter is mainly used for scheduling follicle aspiration, the former is also used for dose adjustments during the COS cycle. Therefore, two different measures of follicular growth were used for quantitative analyses.

**Quantitative analysis of follicular cohort**

In order to quantify follicular growth, the follicles were divided into three categories based on their size, i.e. 10–13, 14–17 and ≥18 mm for each woman. Follicles <10 mm were not included in the analysis. The categories were arbitrarily determined according to the clinical relevance of the follicle sizes. The agreement in the numbers of follicles in each category between the methods was assessed.

The size of the leading follicle becomes more clinically relevant towards the end of the stimulation period and the small variation in the follicle size during the first few days of the COS cycle can lead to a false sense of agreement. Therefore, the comparison of the leading follicle size includes only the subset of women who had a US scan after ≥10 days of gonadotrophin stimulation. The agreement in the size of the leading follicle measured in millimetres by two methods was assessed for this analysis.

**Assessment of agreement between methods**

SonoAVC measurements generated by the investigator who originally captured the ovarian volumes (B.A.) were compared with the 2D manual measurements, which had been used to manage the treatment cycle. Conventional 2D measurements were separately compared with both SonoAVC-generated MFD and $d_v$ measurements. Agreement analyses were done with the Bland–Altman method using Analyse-it® Method Evaluation edition (Analyse-it Software Ltd, Leeds, UK) (Bland and Altman, 1986).

**Assessment of reproducibility of SonoAVC-generated results**

Although SonoAVC automatically generates follicular measurements, a fraction of the SonoAVC measurements require post-processing as
defined above. Therefore, there is a risk that different observers analysing the same raw 3D data set can produce different results. In order to assess the reproducibility of the results, raw ovarian volumes were independently analysed by four different physicians using SonoAVC. Two-way random effects intraclass correlation coefficients (ICCs) with absolute agreement were calculated using SPSS version 17 (IBM, Armonk, NY, USA). In the context of the present study, the ICC is the ratio of variance explained by between-patient variability divided by the total variability, which comprises between-patient variability and between-observer variability. ICC ranges from 0 to 1. The ICC value required for claiming good reproducibility differs across various scales published so far. Most conservative scales require an ICC value $\geq 0.8$ as an indicator of good reproducibility.

Sample size calculation
Currently, there is no standard for sample size calculations used for limit of agreement analyses (Liao, 2009). However, Bland suggests a sample of 100 measurements is adequate in most cases (Bland, 2010).

Results
The mean age of participants was $36.4 \pm 4.6$ years (range 21–45 years), the main indication for IVF was male factor in 24, unexplained infertility in 24, decreased ovarian reserve in 22, polycystic ovarian syndrome in 10, tubal factor in 7 and endometriosis in 7 women. There were four oocyte donors and two women with same-sex partners in the study group. Mean BMI of participants was $25.2 \pm 5.9$ kg/m$^2$ (range 17.7–45.7 kg/m$^2$). Of the participants, 30% had a BMI over 25 kg/m$^2$ and 10.7% had a BMI over 30 kg/m$^2$. There were 39 women who were stimulated with the long GnRH agonist protocol, 35 with the GnRH antagonist protocol and 26 with the microdose flare up protocol. The mean duration of gonadotrophin injections received until the scan day was 10 days (range 5–21 days). There were four women who had a single ovary, five women had endometriomas as diagnosed with 2D US, four women had difficult to visualize ovaries, two women had paraovarian cysts and two women had hydrosalpinges.

The numbers of follicles present in each category are presented in Table I. The agreement between manual measurements and SonoAVC-generated counts based on $d_V$ and MFD is presented in Table II.

When SonoAVC-$d_V$ and manual measurements were compared, the mean difference in the numbers of follicles in the 10–13 mm category was $-1.3$, i.e. the SonoAVC $d_V$-based count was higher than the count by the manual method by 1.3 on average. The mean differences between number of follicles in the 14–17 and $\geq 18$ mm categories were less at $+0.1$ and $+0.5$, respectively, i.e. on average both methods counted almost the same follicle numbers (difference $<1$) in these categories. Similar to the mean differences, the range of the upper and lower limits of agreement ($\pm 2$ SD of the differences) decreased as the follicle sizes increased.

When SonoAVC-MFD and manual measurements were compared, the mean difference in the numbers of follicles in the 10–13 mm category was $-1.1$, i.e. the SonoAVC MFD-based count was higher than the count by the manual method by 1.1 on average. The mean differences between the number of follicles in the 14–17 and $\geq 18$ mm categories were less at $-0.7$ and $-0.3$, respectively, i.e. on average both methods counted almost the same follicle numbers in these
categories (differences < 1). Similar to mean differences, the range of the upper and lower limits of agreement decreased as the follicle sizes increased.

Naturally, the mean differences observed between the methods and particularly the limits of agreement depend on the total number of follicles. In order to assess the agreement between the methods for women with different follicle counts, we divided women into two subgroups based on the mean number of follicle counts in each size category as assessed by 2D scan. The results of these subgroup analyses are presented in Table III. While the mean differences did not change substantially, limits of agreements were wider for women who had more follicles than for those with fewer follicles, as expected.

A post hoc analysis was done to determine whether automated measurements had falsely included endometriomas, paraovarian cysts or hydrosalpinges in follicle counts. All had been correctly identified during the post-processing phase and excluded from follicle counts.

Analyses of the leading follicle size include a subgroup of 48 women who had been scanned after receiving ≥ 10 days of gonadotrophin injections. The mean differences between 2D manual and SonoAVC $d_V$ measurements as well as between 2D manual and SonoAVC MFD measurements were < 1 mm. Full results are presented in Table IV.

Figure 2 provides a visual description of folliculometry results as obtained by 2D manual and two different SonoAVC measurements, based on average follicle counts observed in this study.

The ICC (95% confidence interval) for the number of 10–13, 14–17 and ≥ 18 mm sized follicles were 0.93 (0.90–0.95), 0.90 (0.86–0.92) and 0.79 (0.73–0.84), respectively, for $d_V$-based counts and 0.91 (0.89–0.94), 0.88 (0.84–0.91) and 0.84 (0.79–0.88), respectively, for MFD-based counts. The ICC values for both methods were high across all categories, indicating very good reproducibility.

### Table I. Follicle counts with 2D, SonoAVC $d_V$ and SonoAVC MFD measurements.

<table>
<thead>
<tr>
<th>Measurement technique</th>
<th>Follicle size (mm)</th>
<th>10–13</th>
<th>14–17</th>
<th>≥18</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D US</td>
<td></td>
<td>4.6 ± 3.8 (0–24)</td>
<td>3.3 ± 2.9 (0–13)</td>
<td>1.4 ± 1.7 (0–7)</td>
</tr>
<tr>
<td>SonoAVC MFD</td>
<td></td>
<td>5.7 ± 5.3 (0–32)</td>
<td>4.0 ± 3.5 (0–16)</td>
<td>1.7 ± 1.7 (0–7)</td>
</tr>
<tr>
<td>SonoAVC $d_V$</td>
<td></td>
<td>5.9 ± 5.3 (0–35)</td>
<td>3.4 ± 2.8 (0–12)</td>
<td>0.9 ± 1.2 (0–4)</td>
</tr>
</tbody>
</table>

Values are mean ± SD (range).

### Table II. Agreement in follicle counts with SonoAVC and 2D US.

<table>
<thead>
<tr>
<th>Measurement methods</th>
<th>Follicle size (mm)</th>
<th>10–13</th>
<th>14–17</th>
<th>≥18</th>
</tr>
</thead>
<tbody>
<tr>
<td>SonoAVC $d_V$ versus 2D</td>
<td></td>
<td>-1.3 ± 3.3</td>
<td>0.1 ± 1.7</td>
<td>0.5 ± 1.3</td>
</tr>
<tr>
<td>Mean difference ± SD (2D - $d_V$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper limit of agreement</td>
<td></td>
<td>+5.1</td>
<td>+3.3</td>
<td>+3.0</td>
</tr>
<tr>
<td>Lower limit of agreement</td>
<td></td>
<td>-7.7</td>
<td>-3.5</td>
<td>-1.9</td>
</tr>
<tr>
<td>SonoAVC MFD versus 2D</td>
<td></td>
<td>-1.1 ± 3.1</td>
<td>-0.7 ± 2.0</td>
<td>-0.3 ± 1.2</td>
</tr>
<tr>
<td>Mean difference ± SD (2D - MFD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper limit of agreement</td>
<td></td>
<td>+5.0</td>
<td>+3.2</td>
<td>+2.1</td>
</tr>
<tr>
<td>Lower limit of agreement</td>
<td></td>
<td>-7.2</td>
<td>-4.6</td>
<td>-2.7</td>
</tr>
</tbody>
</table>

### Discussion

It is most unlikely that any two measurement methods will give identical results for all individuals. The relevant parameter is the magnitude of the difference between the new and conventional methods. No statistical significance test exists for such assessment. How far apart the measurements can be without causing problems in clinical practice is a question of judgement (Bland and Altman, 1986). In our opinion, the SonoAVC-generated $d_V$ and MFD measurements were both in good agreement with conventional 2D measurements.

The observed differences between the SonoAVC and manual measurements of follicle counts and the leading follicle size suggest that SonoAVC $d_V$ values tend to be smaller than the conventional 2D measurements. This is evidenced by $d_V$ measurements yielding higher average numbers of follicles in categories of smaller sizes, whereas the average number of follicles in the largest size category is less than that obtained by manual measurement. The cohort seems to be slightly shifted towards smaller sizes with $d_V$ measurements. On the other hand, SonoAVC MFD values tend to be greater than manual measurements and the cohort seems to be slightly shifted towards greater sizes, i.e. upwards in the folliculometry chart, with MFD values. Arguably, these differences can lead to shifting of timing of HCG injection and oocyte collection in some cases, i.e. HCG criteria can be reached earlier with SonoAVC MFD measurements and later with SonoAVC $d_V$ measurements if SonoAVC replaces manual measurements without adjusting for potential differences. However, HCG criteria vary across clinics and clearly there is room for some flexibility (Tan et al., 1992). It is unlikely that the small differences observed between follicle counts and the < 1 mm absolute difference in the size of the leading follicle will affect treatment outcomes. In fact, in a randomized controlled trial comparing the clinical outcomes of COS-IVF cycles in 72 women who were exclusively monitored with either 2D US or SonoAVC, both groups had a similar number of oocytes collected as well as similar pregnancy rates (Rainef-Fenning et al., 2010). The decision-maker who adjusted gonadotrophin dosages and scheduled HCG injection was blinded to the method of measurement. It should be noted that conventional follicular diameter-
A software is substantial. Moreover, the implementation of a user-friendly off-line measurements system requires a sound DICOM infrastructure. What improvements can this costly system offer if it does not improve clinical outcome of treatment? First, SonoAVC saves time during follicle tracking for both the patient and the sonographer. Although post-processing of the initial SonoAVC output and generating the final report can take up to several minutes for some patients, the extent of post-processing required essentially depends on the quality of the grey scale image acquired in the first place. It should be noted that US examination of women with poor quality grey scale images clearly takes more time regardless of the measurement method used. Although we did not record exact times taken for measurements with each method, it is our distinct impression that overall SonoAVC measurements take much less time to complete than manual measurements, especially as one becomes more familiar with the software. This is particularly important for patients with many follicles as it not only saves time but also reduces the discomfort of a prolonged US examination. In fact, previous studies comparing time required for follicle measurements with SonoAVC and with the 2D method have invariably reported substantial time saving provided by SonoAVC (Raine-Fenning et al., 2009a,b). Rodriguez-Fuentes et al. (2010) reported that average time saving reached 7.6 min for women who had ≥10 follicles to measure. Time saving can be further increased by using an off-line system. The ovarian volumes can be captured and transferred to a network by the sonographer, and the patient can leave the examination room without spending any time for measurements. The sonographer can start scanning the next patient while another operator located elsewhere can retrieve the volumes from the network and analyse them with SonoAVC.

Another advantage of SonoAVC is providing a means of standardization of follicle measurements when scans are performed by several sonographers. In busy IVF units such as ours, some patients are scanned by different sonographers in the course of one treatment cycle. Although we are unable to directly compare reproducibility of SonoAVC and 2D measurements due to study design, our results demonstrate very good reproducibility of SonoAVC-generated measurements. ICCs calculated for SonoAVC results were very high, meeting good reproducibility criteria for even conservative scales. In contrast, previous studies assessing reproducibility of 2D

<table>
<thead>
<tr>
<th>Table III Agreement in follicle counts with SonoAVC and 2D US.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement methods</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>SonoAVC d₉ versus 2D</td>
</tr>
<tr>
<td>SonoAVC d₉ versus 2D</td>
</tr>
<tr>
<td>SonoAVC d₉ versus 2D</td>
</tr>
<tr>
<td>SonoAVC MFD versus 2D</td>
</tr>
<tr>
<td>SonoAVC MFD versus 2D</td>
</tr>
<tr>
<td>SonoAVC MFD versus 2D</td>
</tr>
</tbody>
</table>

Based HCG criteria were used for both arms in that study (Raine-Fenning et al., 2010).

The observed trends mentioned above can be explained by the mathematical principles of SonoAVC calculations. Volume-based diameter d₉ is calculated by using a relaxed sphere diameter formula for a perfect sphere having the same volume as the follicle. The diameter of a perfect sphere will always be smaller than the mean of three maximal diameters of an irregular follicle with the same volume regardless of the method of measurement of MFD. On the other hand, SonoAVC can be more accurate than a human observer in identifying the longest diameters in three orthogonal planes, leading to greater MFD values. If conventional 2D scans and SonoAVC are to be used interchangeably in the same IVF unit, perhaps SonoAVC MFD values should be used, as these are calculated with the same methodology (i.e. the mean of the longest diameters in three orthogonal planes), and be rounded down to the next millimetre in order to render SonoAVC-generated MFD values closer to 2D measurement values.

Currently, implementing an automated follicle tracking system requires recent GE Voluson US equipment (V730, E6, E8 or Voluson i) and at least one licence for the 4D view software with the SonoAVC feature. The cost of a single piece of Voluson US equipment with auxiliaries required for conducting 3D scans plus the AVC software is substantial. Moreover, the implementation of a user-friendly off-line measurements system requires a sound DICOM infrastructure. What improvements can this costly system offer if it does not improve clinical outcome of treatment? First, SonoAVC saves time during follicle tracking for both the patient and the sonographer. Although post-processing of the initial SonoAVC output and generating the final report can take up to several minutes for some patients, the extent of post-processing required essentially depends on the quality of the grey scale image acquired in the first place. It should be noted that US examination of women with poor quality grey scale images clearly takes more time regardless of the measurement method used. Although we did not record exact times taken for measurements with each method, it is our distinct impression that overall SonoAVC measurements take much less time to complete than manual measurements, especially as one becomes more familiar with the software. This is particularly important for patients with many follicles as it not only saves time but also reduces the discomfort of a prolonged US examination. In fact, previous studies comparing time required for follicle measurements with SonoAVC and with the 2D method have invariably reported substantial time saving provided by SonoAVC (Raine-Fenning et al., 2009a,b). Rodriguez-Fuentes et al. (2010) reported that average time saving reached 7.6 min for women who had ≥10 follicles to measure. Time saving can be further increased by using an off-line system. The ovarian volumes can be captured and transferred to a network by the sonographer, and the patient can leave the examination room without spending any time for measurements. The sonographer can start scanning the next patient while another operator located elsewhere can retrieve the volumes from the network and analyse them with SonoAVC.

Another advantage of SonoAVC is providing a means of standardization of follicle measurements when scans are performed by several sonographers. In busy IVF units such as ours, some patients are scanned by different sonographers in the course of one treatment cycle. Although we are unable to directly compare reproducibility of SonoAVC and 2D measurements due to study design, our results demonstrate very good reproducibility of SonoAVC-generated measurements. ICCs calculated for SonoAVC results were very high, meeting good reproducibility criteria for even conservative scales. In contrast, previous studies assessing reproducibility of 2D
measurements have shown poor reproducibility (Forman et al., 1991). Another recent study has compared the reproducibility of SonoAVC-generated and 2D measurements, albeit in a monofollicular cycle. Salama et al. (2010) reported higher reproducibility of SonoAVC measurements than that of manual measurements. We believe this is an important advantage particularly for busy IVF clinics.

Thirdly, SonoAVC introduces a chance to implement a quality control system for follicle measurements. It is not practically possible to keep any records of follicle measurements other than the results themselves by the 2D method. If SonoAVC is implemented in routine practice it is possible to save raw ovarian volumes. These volumes are available for retrieval and analysis with SonoAVC at a later time. This enables virtually repeating the US examinations in the absence of the patient should the need arise. Such re-examinations can be periodically repeated for a sample of patients as a method of quality control. Moreover, this gives a chance to monitor results generated by junior staff or trainees in academic units.

Finally, we see three important opportunities that can be provided by SonoAVC. The first is the identification of new follicular volume-based HCG criteria. Volume-based criteria may prove to be a better indicator of oocyte maturity and increase mature oocyte yield. Using follicular volume measured with SonoAVC as the measure of follicular growth combined with volume-based HCG criteria may improve treatment outcome over that achieved with conventional monitoring with follicular diameter and using follicular diameter-based HCG criteria. Secondly, out-of-town patients need to travel several times for follicle monitoring and other procedures for IVF treatment. If they are monitored by local sonographers with less experience in IVF, results can sometimes cause suboptimal management of treatment cycle. SonoAVC can provide an option for these women. If 3D volumes of their ovaries can be captured by local sonographers and transmitted to the IVF centre, they can be analysed with SonoAVC at the treating centre. The more exciting opportunity is that SonoAVC can facilitate self-operated endovaginal telemonitoring. Gerris and De Sutter (2010) have shown that patients can be trained to scan themselves successfully using 2D US. Compared with measuring each and every follicle, capturing the ovarian volume only demands much less skill on the part of the sonographer. It should be easier to teach patients to capture a 3D volume of the ovaries than to teach them to conduct a full follicle count. They can then transmit these volumes to the clinic for assessment with SonoAVC. To take things one step further, if Vscan (GE Healthcare), a new hand-held piece of US equipment, or its equivalents can be developed to capture 3D volumes, this equipment can be given to patients at the beginning of their treatment cycle and they can conduct all monitoring scans in their own time and at the place of their choice (GE Healthcare, 2009).

Authors’ roles
B.A.: conception and design of the study, data acquisition, analysis and interpretation of data, drafting the manuscript, approval of the final version. A.S. and S.L.R.: data acquisition, revising the manuscript, approval of the final version. E.S.-P.: data acquisition, approval of the final version. S.K. and S.L.T.: conception of the study, revising the manuscript, approval of the final version.

Acknowledgement
The authors would like to thank Mrs Najwa Sallman, IT specialist at McGill Reproductive Centre, for her invaluable help in transferring folliculometry results from electronic patient charts to the study database.
References

Bland JM. How can I decide the sample size for a study of agreement between two methods of measurement? http://www-users.york.ac.uk/~mb55/meas/sizemeth.htm (29 May 2010, date last accessed).


