# Intra-observer and inter-observer reproducibility of three-dimensional gray scale and power Doppler ultrasound examinations of the cervix in pregnant women.

L. Rovas, MD, P. Sladkevicius, MD, PhD, E. Strobel, RN, RM, L. Valentin, MD, PhD

Department of Obstetrics and Gynecology, Malmo University Hospital, Malmo, Sweden.

Short title: reproducibility of cervix measurements.

Key words: cervix uteri, three-dimensional imaging, Doppler ultrasound, and reproducibility of results, pregnancy.

Correspondence: Povilas Sladkevicius, Kvinnokliniken, Universitetssjukhuset MAS, S-20502 Malmo, Sweden.

Telephone: +46 40 332094  Fax: +46 40 962600

Email: Povilas.Sladkevicius@obst.mas.lu.se

## ABSTRACT

**Objectives** To determine intra-observer and inter-observer reproducibility of three-dimensional (3D) gray scale and power Doppler ultrasound examinations of the cervix in pregnant women.

**Methods** Thirty-two pregnant women underwent transvaginal 3D gray scale and power Doppler ultrasound examination of the cervix by two examiners. Each observer acquired two volumes, and each observer analysed each of his volumes twice using commercially available software (VOCAL$^{TM}$). The variables analysed were cervical volume ($cm^3$), vascularisation index (VI), flow index (FI) and vascularisation flow index (VFI). Intra-observer repeatability was expressed as the difference between two measurements results (mean difference ± 2SD, i.e., limits of agreement) and as intra-class correlation coefficient (intra-CC). Inter-observer agreement was expressed as the difference between the results of the two observers (limits of agreement) and as inter-class correlation coefficient (inter-CC). The contribution of various factors (examiner, acquisition, analysis of acquired volume) to intra-subject variance was estimated using different analysis of variance models. All statistical analyses were made using log-transformed data. The results presented are those obtained after antilogarithmic transformation, i.e. the results are presented as ratios between two results of the same observer, or as ratios between the results of observer 1 and observer 2.

**Results** All intra-observer and inter-observer log-transformed differences were normally distributed. There was no systematic bias between the two observers.

Both intra- and inter-CC values were high (0.93 - 0.98) for all variables except FI (0.63 - 0.88), despite limits of agreement being wide, especially for VI (up to 0.4 - 2.4) and VFI (up to 0.3 - 2.6). Acquisition explained most of the intra-subject variance of the flow indices, the contribution of examiner and analysis being unimportant.

**Conclusions** Given the wide range between the lower and upper limits of agreement it would probably not be possible to detect anything but large differences or changes in cervical volume or cervical flow indices using current 3D ultrasound technique. Because acquisition explained most of the intra-subject variance, the average of several repeated acquisitions should be used to enhance reproducibility. However, it is not worth doing more than one analysis of an acquired volume, because the effect of analysis on measurement results is small.

## Introduction

Three-dimensional (3D) ultrasound imaging has recently become available for clinical sonographic diagnosis. In theory, 3D ultrasound should provide more accurate volume measurements than conventional two-dimensional (2D) ultrasound[1-3]. One of the latest technical achievements in the field of ultrasonography is 3D imaging combined with power Doppler. It provides a possibility to quantify power Doppler signals in the whole target organ. This is in contrast to 2D ultrasound, where information on vascularisation is restricted to one subjectively chosen 2D plane. 3D power Doppler ultrasonography might be a more appropriate method for semi-quantification of blood supply to an organ than 2D power Doppler ultrasonography. Thus, 3D power Doppler ultrasound may be a suitable tool for studying cervical vascularization during pregnancy, provided that results are reproducible.

The aim of this study was to determine the inter-observer and intra-observer reproducibility of cervical volume calculations and quantification of power Doppler signals in cervical volumes obtained during 3D scanning of the cervix in pregnant women.

## Subjects and methods

The study protocol was approved by the Ethics Committee of the Medical Faculty of Lund University, Sweden. Informed consent was obtained from all participants, after the nature of the procedures had been fully explained.

Pregnant women who voluntarily participated in another research project, which included 3D ultrasound examination of the cervix, were asked to take part. All women were eligible provided that the two ultrasound examiners of the project were available to examine them. Eighteen nulliparous and fourteen multiparous women agreed to participate and were examined as described below**.** Their mean age was 31 years ± 4.2 (standard deviation, SD) range 21-39. They were examined at a mean gestational age of 28 weeks ± 7.3**,** range 17 – 40. None was in labor at the examination.

**Equipment**

The equipment used was a Kretz Voluson 730 ultrasound system (General Electrics, Zipf, Austria) equipped with a 2.8-10 MHz transvaginal transducer. We used a 146° field of view. Identical pre-installed ultrasound settings were used in all women. The Power Doppler settings used were: frequency 3 - 9 MHz, pulse repetition frequency 0.6 kHz, gain -5.0, wall motion filter "low 1".

**Study design**

The study design is shown schematically in Figure 1. All women underwent transvaginal ultrasound examination of the cervix by two experienced examiners (Observer 1, LR and Observer 2, PS). Observer 1 always was the first examiner. Two 3D power Doppler volumes were acquired by each of the two observers. The acquisitions were made at a single examination with approximately twenty seconds between the first and second acquisition of each observer and only a few minutes between the examination of the first and second observer. Total examination time was approximately 10 min. The observers were not present during each other's examinations or calculations and were kept unaware of each other's results until all acquired volumes had been analysed. Analysis of stored volumes was done off-line, each observer analysing each of his own acquired volumes twice**.** In addition, Observer 1 also once analysed the second volume obtained by Observer 2.

The women were examined in the lithotomy position with an empty bladder. The ultrasound probe was slowly introduced into the vagina and care was taken to avoid undue pressure on the cervix. After a satisfactory gray scale image of the cervix had been obtained, the probe was withdrawn until the image became blurred. Then the probe was gradually advanced again with only enough pressure to restore a satisfactory image. A sagittal view of the cervix where the internal os, the cervical canal and the external os were all seen at the same time was obtained (Figure 2a, upper left quadrant).

Then the system was switched into the power Doppler mode and then into the 3D mode. The cervix was centralized within the 3D sector appearing on the ultrasound screen, and data were obtained by holding the transducer stationary while its crystals were mechanically rotated across the sector with a sweep angle of 90°. The fast volume acquisition (low resolution) setting was always used to minimize periodic flashing artefacts arising from pulsation of the uterine arteries and from fetal movements. The duration of the volume acquisition was 15 – 20 s depending on the dimension of the 3D sector. The scanned volumes were stored digitally for analysis off-line.

Analyses of the stored cervical volumes were done off line. Cervical volume ($cm^3$) and power Doppler flow indices were calculated using the Virtual Organ Computer-aided AnaLysis (VOCAL$^{TM}$), which is integrated into the Voluson 730 ultrasound system. The following Doppler indices were calculated: vascularization index (VI), flow index (FI) and vascularization flow index (VFI). VI is the ratio between the color voxels and the total number of voxels in the volume. It reflects the percentage of the volume consisting of blood vessels. FI is calculated as the sum of weighted color voxels divided by the number of color voxels. It reflects the mean energy per color voxel reflected from the blood corpuscles in the vessels of the volume, i.e., the more blood corpuscles the higher the FI values. VFI is the sum of weighted color voxels divided by the total number of voxels. It reflects both the proportion of tissue consisting of vessels and the number of blood corpuscles in the blood vessels[4,5] .

The acquired volumes were manipulated to obtain reformatted multiplanar views of the cervix in the mid-sagittal, axial and coronal planes. All cervical measurements were performed on these multiplanar images, and results were documented on hard copies**.** The contour mode in the VOCAL$^{TM}$ program was set to manual. The longitudinal view was used as the reference image. The rotation steps were 30° i.e., six contours of the cervix were drawn manually using the roller ball cursor of the system (Figure 2a). Care was taken not to include

the lower uterine segment, the vaginal wall, and particularly the large uterine arteries when drawing the contours (Figure 2a). The hyperechogenic or hypoechogenic line between the cervix and the vaginal wall (Figure 2a), and the internal and external cervical os were used as landmarks. Once all contours had been drawn, the volume and power Doppler flow indices of the cervix were computed automatically (Figure 2b). Before the start of our study the two ultrasound examiners had practiced volume calculations together and agreed upon which landmarks to use when drawing the contours of the cervix.

**Statistical analysis**

Intra-observer repeatability was expressed as the difference between two measurement results obtained by the same observer. The difference between the first and second analysis of the first acquired volume was calculated as well as that between the first analysis of the first volume and the first analysis of the second volume (Figure 1). The difference between the measured values was plotted against the mean of the two measurements to assess the relationship between the difference and the magnitude of the measurements. Limits of agreement (mean difference ± 2 SD) were calculated as described by Bland and Altman[6,7]. Systematic bias between the first and second analysis was determined by calculating the 95% confidence interval (CI) for the mean difference (mean difference $\underline{+}$ 2 standard errors, SE). If zero lay within this interval, no bias was assumed to exist. Intra-observer repeatability was also expressed as the intra-class correlation coefficient (intra-CC), variance components being estimated from different analysis of variance models[8].

In all calculations used for determining inter-observer differences one measurement value per observer was used. These measurement values were: the mean of all four measurement results, the result of the first analysis of the first volume, the mean of the first analyses of the first and second volume, and the first analysis of each observer of the second volume obtained by Observer 2 (Figure 1). To assess systematic bias between the two observers, and to assess

the relationship between the difference between their measured values and the magnitude of the measured values, the differences between the measurements of the two observers were plotted against the means of the measurements obtained by both observers[6]. Bias between the two observers and limits of agreement were calculated as described above for intra-observer reproducibility. Inter-observer agreement was also expressed as the inter-class correlation coefficient (Inter-CC)[8], variance components being estimated by analysis of variance as described above.

Because all intra-observer differences increased with the magnitude of the measurements values, the values were subjected to logarithmic transformation, whereupon the correlation disappeared[6]**.** Inter-observer differences for VI and VFI also increased with increasing magnitude of the measurements values. When these values were submitted to logarithmic transformation the correlation disappeared. There was no clear tendency for inter-observer differences in FI or cervical volume to increase with increasing magnitude of the measurement values, but when the values for FI and volume were logarithmically transformed, there was no tendency for a correlation between the differences and the magnitude of the measurement values. Therefore, all statistical analyses (including calculation of intra- and inter-class correlation coefficients) were made using log-transformed data. The results presented are those obtained after antilogarithmic transformation, i.e., the results describing intra-observer differences are presented as ratios between two measurement results of the same observer, and results describing inter-observer differences are presented as ratios between the results of observer PS and observer LR[6].

Statistical calculations were made using StatView®, version 5 (SAS Institute Inc., USA, 1999) and the Statistical Package for the Social Sciences software, version 10.0.5 (Chicago, IL, USA, 1999).

**Results**

Measurement results (absolute values and log-transformed values) are shown in Table 1. The log-transformed values were normally distributed.

Intra-observer reproducibility is shown in Tables 2, 3 and 4. All intra-observer differences were normally distributed. There was no systematic bias between any paired measurements results.

Inter-observer reproducibility is shown in Tables 5, 6, 7, 8 and 9. All inter-observer differences were normally distributed. There was no systematic bias between the two observers.

Both intra- and inter-CC values were high for all variables except FI. Limits of agreement were wide, especially for VI and VFI. Acquisition explained most of the intra-subject variance in flow indices.

**Discussion**

The intra-CC and inter-CC indicate the proportion of the total variance in measurement results that can be explained by differences between the individuals examined. A high intra-CC or inter-CC indicates that the measurements can be used to discriminate between individuals. Values for inter-CC and intra-CC from 0.75 to 1.0 are said to be acceptable[9]. The more variable is the population investigated, the greater are the intra-CC and inter-CC, and the less variable the population is, the smaller are the intra-CC and inter-CC**.** Therefore, not only should the intra-CC and inter-CC be used to indicate agreement, but also the absolute variance (or standard deviation) of the differences should be taken into account[7].

The high intra- and inter-CC values in our study reflect the substantial variability of our study population. The wide limits of agreement show that our measurement results were not precise, in particular the Doppler results were imprecise. Both in our study and in a study

assessing the reproducibility of 3D power Doppler measurements in ovaries[4] intra-CC and inter-CC values were lowest for FI.

Why are 3D power Doppler indices difficult to reproduce? Factors likely to affect the results are: pressure on the cervix with the transducer during scanning, uterine contractions (at least, these are known to have an effect on blood flow velocities in the uterine arteries[10,11], vena cava syndrome (even thought, no woman in our study had vena cava syndrome during the ultrasound examination), and physiological changes in cervical blood circulation during the examination. Technical factors are probably very important, e.g., whether an ultrasound beam hits a vessel during systole or diastole. Even though power Doppler ultrasound is said to be angle independent [12], this is not entirely true. An ultrasound beam hitting the blood stream in a vessel under a 90º angle will generate no Doppler shift. Thus small changes in the position of the vaginal transducer may result in differences in the power Doppler signals in the volume acquired. Moreover, there are difficulties with defining the border between the cervix and lower uterine segment and between the cervix and the vagina. These difficulties were also emphasised by another research team who studied cervical volume using 3D ultrasound[13]. However, after some training it is possible to get satisfactory results, as illustrated by the reasonably acceptable reproducibility of cervical volume calculations in our study.

Given the wide range between the upper and lower limits of agreement, it would probably be impossible to detect small true intra-individual changes in cervical volume or cervical flow indices in a longitudinal study where replicate measurements were to be taken by the same observer or by different observers. It would also almost certainly be impossible to detect anything but large differences between two or more study populations. The magnitude of clinically important changes or differences in cervical volume or cervical flow indices is currently unknown. Because acquisition contributed most to the intra-subject variance, the

average of several repeated acquisitions should be used to enhance measurement reproducibility. However, it is not worth doing more than one analysis of an acquired volume, because the effect of analysis on measurement results is small.

**Acknowledgments**

This study was supported by the Swedish Medical Research Council (grants nos. K2001-72X-11605-06A and K2002-72X-11605-07B), two governmental grants (Landstingsfinansierad regional forskning, Region Skåne, and ALF-medel), and Funds administered by Malmö University Hospital.

**References**

1. Bega G, Lev-Toaff A, Kuhlman K, Berghella V, Parker L, Goldberg B, Wapner R. Three-dimensional multiplanar transvaginal ultrasound of the cervix in pregnancy. *Ultrasound Obstet Gynecol* 2000; **16**: 351-358.

2. Chou CY, Hsu KF, Wan ST, Huang SC, Tzeng CC, Huang KE. Accuracy of three-dimensional ultrasonography in volume estimation of cervical carcinoma. *Gynecol Oncol* 1997; **66**: 89-93.

3. Tom Farrell, Leslie JR, Chien PFW, Augustsson P. The reability and validity of three dimensional ultrasound volumetric measurements using an in vitro ballon and in vivo uterine model. *Br J Obstet Gynecol* 2001; **108**: 573-582.

4. Järvelä Y, Sladkevicius P, Tekay AH, Campbell S, Nargund G. Intraobserver and interobserver variability of ovarian volume, gray-scale and color flow indices obtained using transvaginal three-dimensional power Doppler ultrasonography. *Ultrasound Obstet Gynecol* 2003; **21**: 277-282.

5. Pairleitner H, Steiner H, Hasenoehrl G, Staudach A. Three-dimensional power Doppler sonography: imaging and quantifying blood flow and vascularization. *Ultrasound Obstet Gynecol*. 1999; **14**:139-143.

6. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307-310.

7. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Meth Med Res* 1999; **8**: 135-160.

8. Scherjon SA, Kok JH, Oosting H, Zondervan HA. Intra-observer and inter-observer reliability of the pulsatility index calculated from pulsed Doppler flow velocity waveforms in three fetal vessels. *Br J Obstet Gynecol* 1993; **100**:134-138.

9. Kramer MS, Feinstein AR. Clinical biostatistics. LIV. The biostatistics of concordance. *Clin Pharmacol Ther* 1981; **29**: 111-123.

10. Valentin L, Sladkevicius P, Kindahl H, Broeders A, Marsel K, Melin P. Effects of a Vasopressin antagonist in women with dysmenorrhoea. *Gynecol and Obstetr Invest* 2000; **50**:170-177.

11. Li H, Gudmundsson S, Olofsson P. Uterine artery blood flow velocity waveforms during uterine contractions. *Ultrasound Obstet Gynecol*. 2003; **22**: 578-585.

12. Martinoli C, Derchi LE, Rizzatto G, Solbiati L. Power Doppler sonography: general principles, clinical applications, and future prospects. *Eur Radiol* 1998; **8**: 1224-1235.

13. Hoesli IM, Surbek DV, Tercanli S, Holzgreve W. Three dimensional volume measurement of the cervix during pregnancy compared to conventional 2D-sonography. *Int J Gynaecol Obstet*. 1999; **64**: 115-119.

**Figure1**

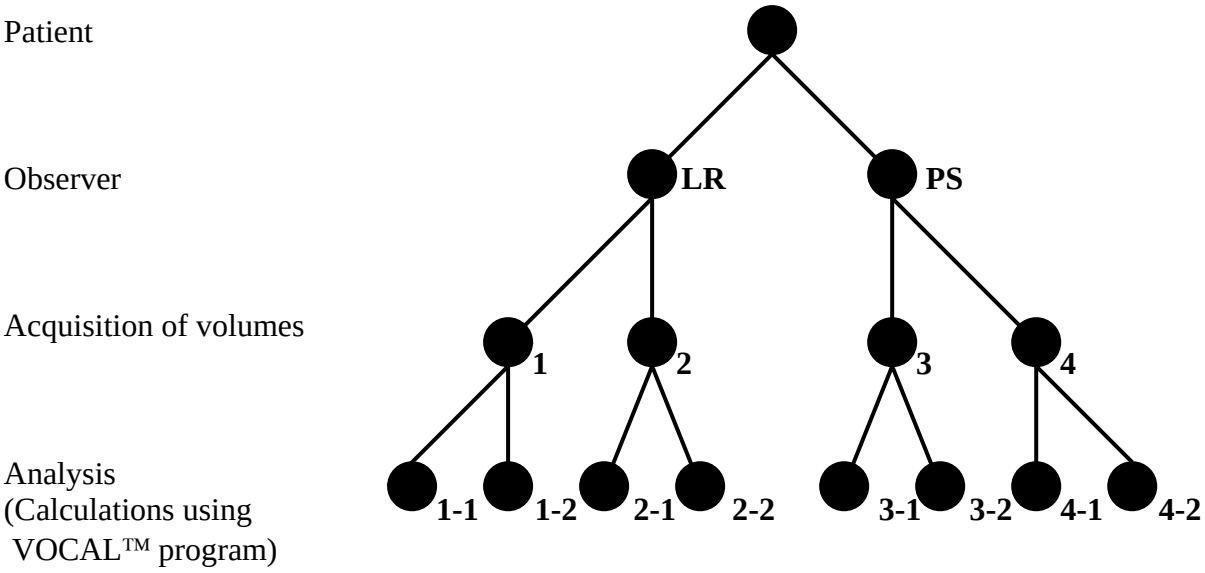

Patient
Observer
Acquisition of volumes
Analysis
(Calculations using
VOCAL™ program)
LR
PS
1
2
3
4
1-1
1-2
2-1
2-2
3-1
3-2
4-1
4-2

**Figure 2a**

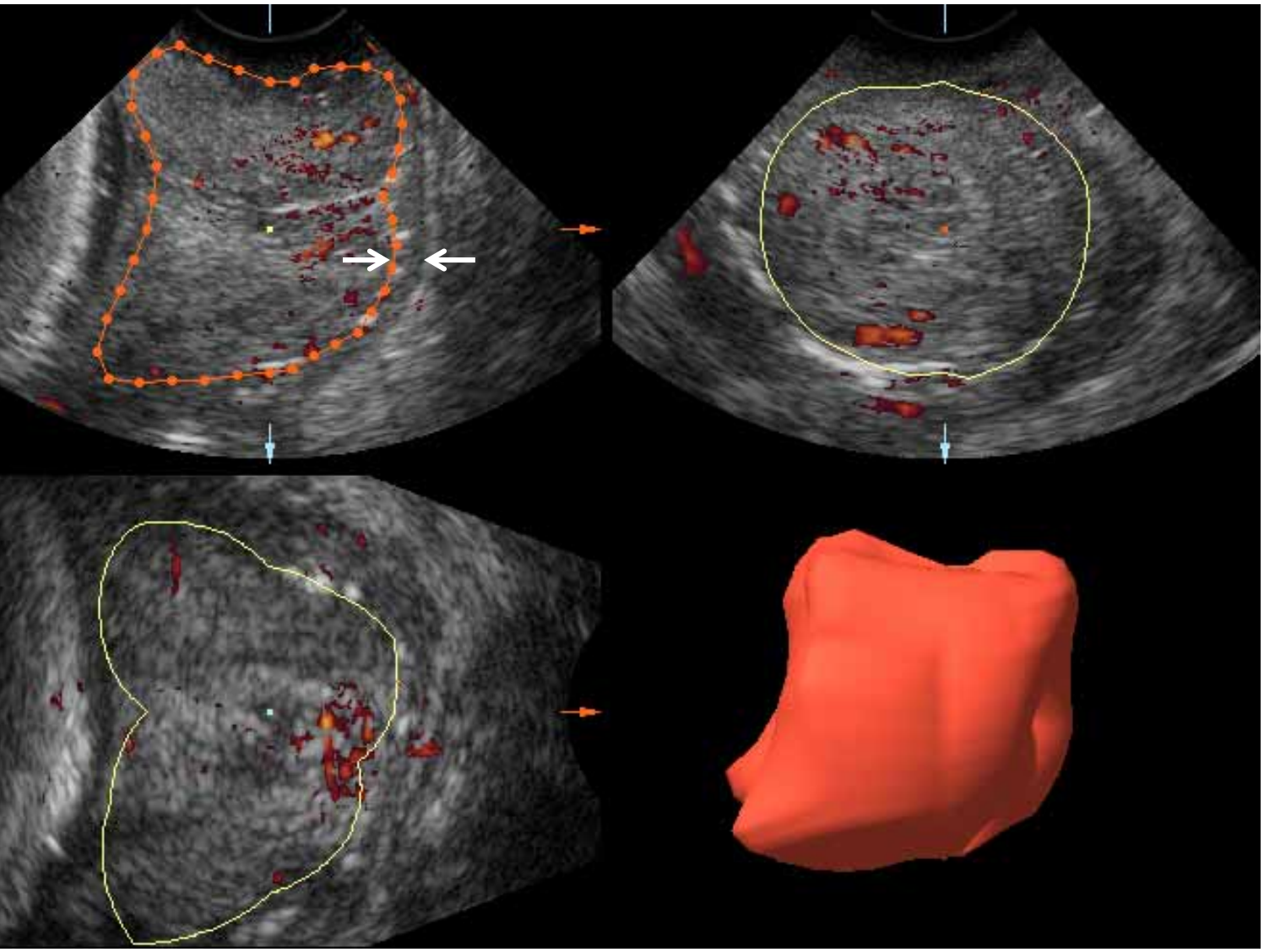
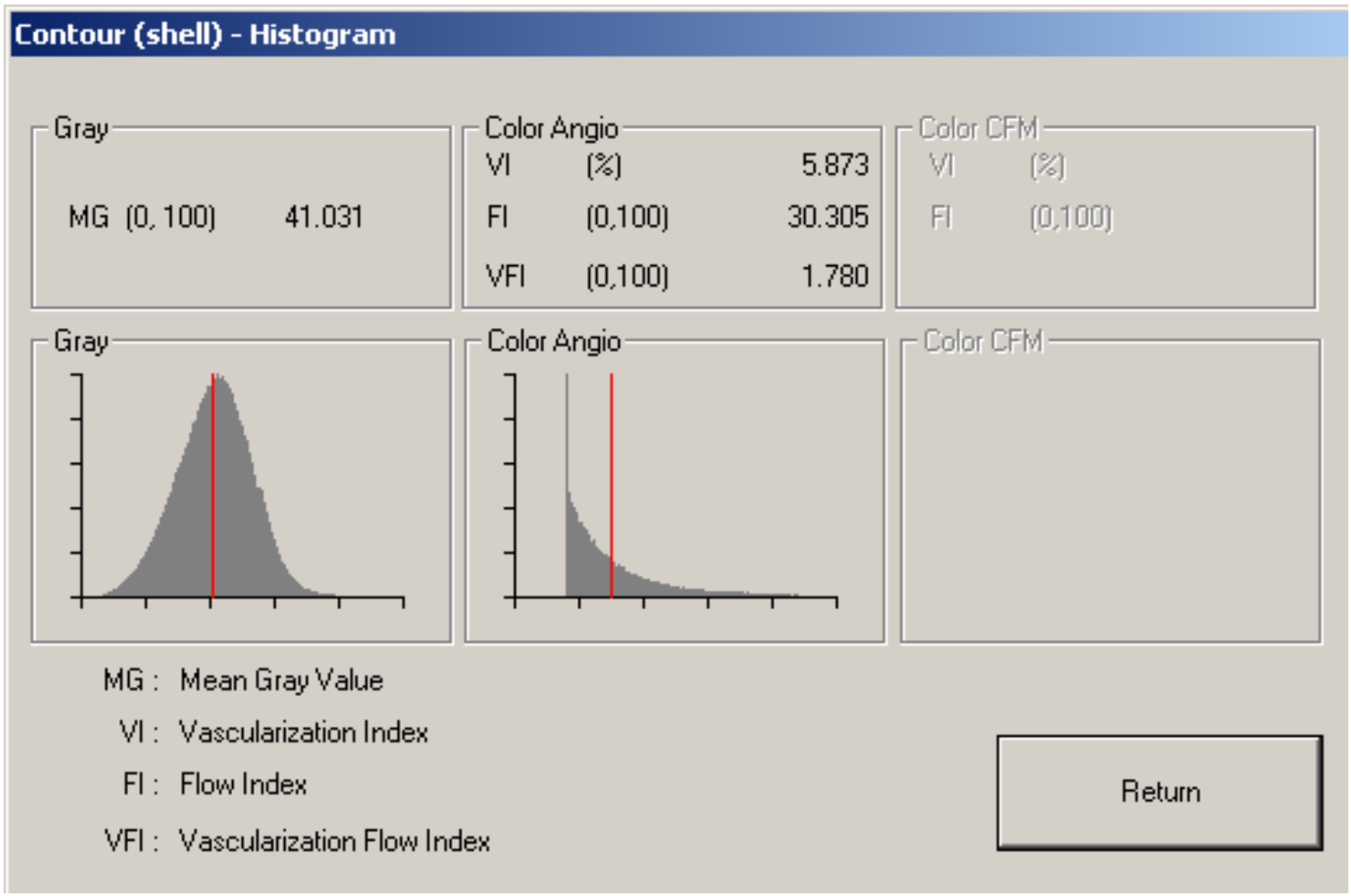


**Figure 2b**


Contour (shell) - Histogram
Gray
MG (0, 100) 41.031
Color Angio
VI (%) 5.873
FI (0,100) 30.305
VFI (0,100) 1.780
Color CFM
VI (%)
FI (0,100)
Gray
Color Angio
Color CFM
MG : Mean Gray Value
VI : Vascularization Index
FI : Flow Index
VFI : Vascularization Flow Index
Return

**Legends:**

**Figure 1** Schematic drawing of the study design. To determine intra-observer reproducibility, the following differences were calculated: 1-1 minus 1-2 and 3-1 minus 3-2 (Table 2); 1-1 minus 2-1 and 3-1 minus 4-1 (Table 3). To determine inter-observer reproducibility the following differences were calculated: 1-1 minus 3-1 (Table 5); (1-1 + 2-1)/2 minus (3-1 + 4-1)/2 (Table 6); (1-1 + 1-2 + 2-1 + 2-2)/4 minus (3-1 + 3-2 + 4-1 + 4-2)/4 (Table 7); In addition observer LR also analysed 4-1 (Table 8). **VOCAL™ is the commercial software used for calculation of volume and flow indices.**

**Figure 2** Three-dimensional ultrasound measurement of cervical volume. (a) multiplanar display of the cervix: longitudinal plane in the upper left quadrant, transverse plane in the upper right quadrant and coronal plane in the lower left quadrant. The resultant three-dimensional model can be seen in the lower right image. The tracing of the cervix is demarcated by lines. **The two white arrows demarcate the thickness of the vaginal wall.** (b) Vascular indices as shown on the ultrasound screen.

**Table 1** Results of cervical measurements obtained by each observer

| Observer Parameter | Mean | SD | Median | Range |
|---|---|---|---|---|
| Absolute values | | | | |
| LR volume ($cm^3$) | 38.39 | 14.61 | 39.05 | 18.0 - 85.60 |
| PS volume ($cm^3$) | 37.47 | 14.82 | 36.95 | 13.70 - 84.00 |
| LR VI | 5.10 | 4.66 | 2.90 | 0.30 - 17.60 |
| PS VI | 4.57 | 4.04 | 3.20 | 0.40 - 16.40 |
| LR FI | 31.04 | 3.83 | 30.80 | 24.00 - 43.90 |
| PS FI | 30.22 | 3.32 | 29.95 | 22.50 - 39.60 |
| LR VFI | 1.61 | 1.57 | 0.80 | 0.10 - 5.70 |
| PS VFI | 1.39 | 1.32 | 0.90 | 0.10 - 6.00 |
| Logged values | | | | |
| LR volume ($cm^3$) | 1.554 | 0.163 | 1.592 | 1.255 - 1.932 |
| PS volume ($cm^3$) | 1.541 | 0.170 | 1.568 | 1.137 - 1.924 |
| LR VI | 0.513 | 0.438 | 0.462 | -0.523 - 1.246 |
| PS VI | 0.488 | 0.407 | 0.505 | -0.398 - 1.215 |
| LR FI | 1.489 | 0.052 | 1.489 | 1.380 - 1.642 |
| PS FI | 1.478 | 0.048 | 1.476 | 1.352 - 1.598 |
| LR VFI | -0.023 | 0.481 | -0.097 | -1.000 - 0.758 |
| PS VFI | -0.059 | 0.450 | -0.046 | -1.000 - 0.778 |

LR - first observer; PS - second observer; VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation. The results presented are based on all 128 values.

**Table 2** Intra-observer differences between the first and second analysis of the first volume acquired. Antilogged values corresponding to ratios between the first and second analysis are shown.

| | | Inter - observer difference (expressed as a ratio) | | | |
|---|---|---|---|---|---|
| Observer | Parameter | Mean | Limits of agreement | 95% CI | Intra-CC |
| LR | volume, $cm^3$ | 0.991 | 0.875 – 1.122 | 0.968 – 1.006 | 0.99 |
| PS | volume $cm^3$ | 0.995 | 0.836 – 1.186 | 0.964 – 1.028 | 0.97 |
| LR | VI | 0.986 | 0.787 – 1.236 | 0.946 – 1.028 | 0.99 |
| PS | VI | 0.986 | 0.787 – 1.236 | 0.946 – 1.028 | 0.99 |
| LR | FI | 0.991 | 0.916 – 1.072 | 0.977 – 1.005 | 0.96 |
| PS | FI | 1.007 | 0.939 – 1.097 | 0.993 – 1.021 | 0.95 |
| LR | VFI | 0.984 | 0.746 – 1.297 | 0.935 – 1.035 | 0.99 |
| PS | VFI | 0.986 | 0.728 – 1.337 | 0.933 – 1.042 | 0.99 |

LR - first observer; PS - second observer; VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; CI, confidence interval; Intra-CC, intra-class correlation coefficient.

**Table 3** Intra-observer differences between results of the analyses of the first and second volume acquired. Antilogged values corresponding to ratios between the results of the first and second volume are shown.

| | | Intra – observer difference (expressed as a ratio) | | | |
|---|---|---|---|---|---|
| Observer | Parameter | Mean | Limits of agreement | 95% CI | Intra-CC |
| LR | volume, $cm^3$ | 1.021 | 0.873 – 1.194 | 0.993 – 1.050 | 0.98 |
| PS | volume $cm^3$ | 0.986 | 0.776 – 1.253 | 0.946 – 1.028 | 0.96 |
| LR | VI | 1.014 | 0.503 – 2.042 | 0.895 – 1.148 | 0.94 |
| PS | VI | 1.067 | 0.472 – 2.410 | 0.925 – 1.230 | 0.91 |
| LR | FI | 1.019 | 0.836 – 1.242 | 0.982 – 1.057 | 0.66 |
| PS | FI | 0.993 | 0.818 – 1.205 | 0.962 – 1.026 | 0.63 |
| LR | VFI | 1.045 | 0.436 – 2.506 | 0.893 – 1.222 | 0.93 |
| PS | VFI | 1.079 | 0.438 – 2.661 | 0.918 – 1.268 | 0.96 |

LR - first observer; PS - second observer; VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; CI, confidence interval; Intra-CC, intra-class correlation coefficient.

**Table 4** Contribution of various factors to intra-observer variance in measurement results

| | Volume | VI | FI | VFI |
|---|---|---|---|---|
| Contribution (%) to variance for observer LR | | | | |
| Patient | 95.7 | 94.3 | 66.4 | 92.6 |
| Acquisition of volume | 2.0 | 5.2 | 26.6 | 6.5 |
| Analysis of volume | 2.3 | 0.5 | 6.9 | 0.9 |
| Contribution (%) to variance for observer PS | | | | |
| Patient | 98.1 | 90.4 | 52.0 | 88.1 |
| Acquisition of volume | 0.4 | 8.1 | 39.4 | 9.3 |
| Analysis of volume | 1.5 | 1.5 | 8.6 | 2.6 |

**Table 5** Inter-observer differences between the results of the first analysis of the first volume calculated by each observer. Antilogged values are shown corresponding to a ratio between the results of the first and second observer.

| | Inter – observer difference (expressed as a ratio) | | | |
|---|---|---|---|---|
| Parameter | Mean | Limits of agreement | 95% CI | Inter-CC |
| Volume | 0.962 | 0.713 – 1.297 | 0.914 – 1.012 | 0.92 |
| VI | 0.959 | 0.377 – 2.443 | 0.813 – 1.132 | 0.89 |
| FI | 0.966 | 0.785 – 1.189 | 0.931 – 1.002 | 0.59 |
| VFI | 0.931 | 0.332 – 2.612 | 0.774 – 1.119 | 0.89 |

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; CI, confidence interval; Inter-CC, inter-class correlation coefficient.

**Table 6** Inter-observer differences between the means of two measurements per observer (first analysis of each volume). Antilogged values are shown corresponding to a ratio between the results of the first and second observer.

| | Inter – observer difference (expressed as a ratio) | | | |
|---|---|---|---|---|
| Parameter | Mean | Limits of agreement | 95% CI | Inter-CC |
| Volume | 0.977 | 0.759 – 1.259 | 0.933 – 1.023 | 0.95 |
| VI | 0.938 | 0.391 – 2.249 | 0.802 – 1.096 | 0.90 |
| FI | 0.977 | 0.832 – 1.148 | 0.951 – 1.005 | 0.70 |
| VFI | 0.916 | 0.342 – 2.455 | 0.769 – 1.091 | 0.90 |

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; CI, confidence interval; Inter-CC, inter-class correlation coefficient.

**Table 7** Inter-observer differences between the means of four measurements per observer. Antilogged values corresponding a ratio between the results of the first and second observer are shown.

| | Inter – observer difference (expressed as a ratio) | | | |
|---|---|---|---|---|
| Parameter | Mean | Limits of agreement | 95% CI | Inter-CC |
| Volume | 0.971 | 0.782 – 1.205 | 0.935 – 1.007 | 0.96 |
| VI | 0.944 | 0.385 – 2.317 | 0.804 – 1.109 | 0.89 |
| FI | 0.974 | 0.841 – 1.135 | 0.927 – 1.002 | 0.71 |
| VFI | 0.923 | 0.346 – 2.460 | 0.774 – 1.099 | 0.89 |

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; CI, confidence interval; Inter-CC, inter-class correlation coefficient.

**Table 8** Inter-observer differences between the first and second observer when they analysed the same volume. Antilogged values are shown corresponding to a ratio between the first and second observer.

| | Inter – observer difference (expressed as a ratio) | | | |
|---|---|---|---|---|
| Parameter | Mean | Limits of agreement | 95% CI | Inter-CC |
| Volume | 0.974 | 0.833 – 1.140 | 0.948 – 1.002 | 0.92 |
| VI | 1.000 | 0.734 – 1.361 | 0.946 – 1.056 | 0.98 |
| FI | 1.006 | 0.918 – 1.104 | 0.993 – 1.020 | 0.88 |
| VFI | 1.002 | 0.632 – 1.588 | 0.922 – 1.088 | 0.95 |

VI, vascularisation index; FI, flow index; VFI, vascularisation flow index; SD, standard deviation; CI, confidence interval; Inter-CC, inter-class correlations coefficient.

**Table 9** Contribution of various factors to inter-observer variance in measurement results

| | Volume | VI | FI | VFI |
|---|---|---|---|---|
| Contribution (%) to variance based on eight measurement results per woman | | | | |
| Patient | 94.8 | 88.0 | 57.6 | 86.5 |
| Observer | 0.2 | 0.1 | 1.8 | 0.1 |
| Acquisition | 3.1 | 11.0 | 33.0 | 11.6 |
| Analysis | 1.9 | 0.9 | 7.6 | 1.8 |