Quality Assurance in Ultrasonographic Measurements of the Myomatous Uterus

Päivi H. Rovio[*,1], Tiina Luukkaala[2], Maarit Vuento[1], Sinikka Oksa[1] and Pentti K. Heinonen[1,3]

[1]*Department of Obstetrics and Gynecology, Tampere University Hospital, Tampere, Finland*

[2]*Tampere School of Public Health, University of Tampere, and Research Unit, Pirkanmaa Hospital District, Tampere, Finland*

[3]*Medical School, University of Tampere, Tampere, Finland*

Abstract: It is important to evaluate quality control in health care. The aim of the current study was to evaluate different sources of variation in estimating the preoperative weight of the myomatous uterus by using repeatability and reproducibility (R&R) method. Material and Methods: Estimation of the total weight of the uterus by a formula combining the formulas for the prolate ellipsoid and cylinder was preoperatively determined using a transvaginal ultrasound probe. Three physicians repeated each round of measurements three times, producing in total 108 findings (12 subjects x 3 investigators x 3 rounds) in 12 women with symptomatic leiomyomas scheduled to undergo hysterectomy. Variation was divided into different components: physicians, patients, and repeated measurements. Variation due to differences across repeated measurements (repeatability), across physicians (reproducibility), and across patients (variability) was then estimated. The estimates of uterine weights were compared to the true weight of the hysterectomy specimen. Results: The more experienced the physician was in taking the ultrasound measurements the less deviation was observed between her own three measurements. Repeatability was 28 %, reproducibility 0% and patient-to-patient variation 72%. There was no significant difference between the accuracy of the measurements of the three physicians. Conclusions: The experience of the physician had an effect on repeatability but not on reproducibility in estimating uterine weight by ultrasound. Our results indicate the importance of conscientiousness in taking measurements of the uterus and cervix.

Keywords: Quality control, ultrasonography, uterine leiomyoma, uterine weight.

INTRODUCTION

The size of the uterus has clinical significance in cases of uterine malignancy, in following up growth of myomas or when adenomyosis is suspected. It is also an important factor in assessing the hysterectomy route most appropriate for the individual patient [1]. Myomas enlarge the uterus and are the most common reason for a hysterectomy [2]. It is common gynecologic practice to estimate the size of nongravid enlarged uteri by bimanual examination and making a comparison with pregnant uteri of comparable size. This procedure provides a rough guide about true size of the uterus [3-4]. Ultrasonographic estimation of the volume and weight of the uterus has been introduced by a number of clinicians [3-7]. We have recently applied a new combined formula where the volume of the uterine corpus and cervix are measured separately by transvaginal ultrasound probe [8]. This proved to be a more accurate means of estimating the true volume of the total uterus than using the traditional method.

It is important to evaluate quality control in health care. This is a process which seeks to monitor and assess the actual quality of care given to an individual patient, to a certain patient group, or to a population. When analysing and developing the quality of treatment, it is important to ensure low overall variation in the treatment process and to ascertain how this variation can be focused into different components within that process. Repeatability and reproducibility (R&R) method is widely applied in the analysis of industrial processes [9], but rarely in quality control in medicine [10].

We carried out a prospective study using R&R method to find out whether the experience of the physician has an influence on the preoperative estimation of the volume of the myomatous uterus and whether there is variation across a single physician´s own measurements.

MATERIALS AND METHODOLOGY

Patients

The study group comprised 12 patients with symptomatic uterine leiomyomas scheduled for a hysterectomy at the Department of Obstetrics and Gynaecology, University Hospital of Tampere, Finland. Their age ranged from 42 to 56 years (median age 49.5 years) and three of patients were postmenopausal. Four used hormone replacement therapy. Body mass index ranged from 22.7 to 35.8 kg/m² (median 26.7 kg/m²) and four of the patients were nulliparous. All the subjects gave written informed consent and the study protocol was approved by the Ethics Committee of the University Hospital of Tampere.

*Address correspondence to this author at the Department of Obstetrics and Gynecology, FI-33521, Tampere University Hospital, Finland; Tel: +358 3 3116 7682; Fax: +358 3 3116 4360; E-mail: paivi.rovio@pshp.fi

1874-3471/08

Ultrasound Measurements

Preoperative pelvic images were taken of each subject within one week prior to hysterectomy by three physicians. One of these (physician 1) was a highly experienced senior gynecologist with oncological subspeciality. The other two physicians were resident postgraduates training in Obstetrics and Gynecology. Physician 2 was experienced in measuring uterine dimensions by vaginal and abdominal ultrasound, and the other (physician 3) was not. An ultrasound sector scanner (Nemio 20®, Toshiba Co., Tochigi-Ken, Japan) with a 6 MHz vaginal probe was used for all the imaging, with the subject supine and with an empty bladder. If the field of view of the vaginal probe was not wide enough to measure the dimensions of a large uterus, a 3.75 MHz abdominal probe was used instead. The dimensions of the uterine corpus and cervix were measured as previously described [8]. Each physician determined all the measurements independently and repeated each round of measurements three times during the scanning of each patient. Thus the total number of findings was 108 (12 subjects x 3 investigators x 3 rounds).

Calculation of Volume

Calculation of the volume of the uterine corpus was done according to a geometric formula for a prolate ellipsoid based on the length (L), width (W) and anteroposterior diameter (AP) of the corpus: volume = 0.5236 x L x W x AP. Calculation of the volume of the cervix was done according to a geometric formula for a cylinder based on its anteroposterior diameter (D) and length (CL): volume = $\pi (D/2)^2$ x CL. The volume of the corpus and that of the cervix were summed and the total volume obtained.

Measurements of the Preparates

At hysterectomy the uterine corpus and cervix were immediately measured after removal (physician 2). Fallopian tubes and ovaries were carefully removed from the uterus and the cervix cut from the corpus before measurements. The three maximal dimensions of the length, depth and width of the corpus and the length and diameter of the cervix were obtained using pair of compasses and a centimetre scale on the same axis as that described for the ultrasonographic measurements. The weights of the uterine corpus and cervix were measured with a digital scale and summed.

Statistics

In the present study the repeatability and reproducibility (R&R) method was applied. The design of the study was experimental, the aim being estimate different sources of variation in the size of uterus by three physicians repeating the same measurements three times. Because the distribution of the difference between real and calculated weight was skewed, the range method was used.

According to the standard Gage R&R terminology physicians stand for operators, patients for parts and repeated differences between real weight and calculated weight stand for trials. In statistical terms the following variance components were estimated: repeatability (difference across measurements), reproducibility (difference across physicians) and variability (difference across patients). In other words, repeatability describes intraphysician variation, i.e., how a given physician repeats the same planning process. Reproducibility describes interphysician variation, i.e., how different physicians follow the same planning process, and variability describes interpatient variation, i.e. how well the same physician can repeat the planning process for different kinds of patients. The total error - also known as the combined R&R effect - includes repeatability and reproducibility, and only patient-to-patient variation is excluded.

The results are presented in graphical form as range charts and summary box plots. The Repeatability & Reproducibility Summary Plot shows the individual average deviation for each physician. In the plot each trial is presented as a dot and each physician´s trials are connected by a vertical line. In the range charts the warning limit is the so-called three times standard deviation (99 %) line. Observations above that limit are regarded as outliers. In the ideal case, no variation will be found due to trials or due to physicians; the only variation that exists will be due to patients. The present statistical analyses were performed by Statistica/W (Version 5.1, 98 edition, Statsoft. Inc, Tulsa, OK, USA).

RESULTS

The weight of the total removed uterus (corpus and cervix) ranged from 106.9 to 1097.4 g (median 213.9 g). The weight of the removed uterine corpus ranged from 75.7 to 1074.7 g (median 181.8 g) and that of the uterine cervix from 20.5 to 44.5 g (median 30.2 g). Use of a transabdominal ultrasound probe was partially necessary in two patients with the largest uteruses.

Due to combined R&R the difference error between real uterus weight and uterus weight measured by ultrasound was 28 % consisting merely of repeatability (variation across measurements). This means that physician failed to get a completely same measurement results from each three measuring round. The experience of the physician had no effect since variation due to physicians was 0 %. Patient-to-patient variation, which was 72 %, was the largest (Table **1**).

Table 1. Estimated Components of Variance in Measuring Difference Between Real Uterus Weight and Measured Weight (Number of Physicians 3, Patients 12, Measurements 3)

Source (Expected MS)	Estimated SD	Estimated Variance	% of R & R	% of Total
Repeatability	38.46	1478.91	100.00	27.96
Reproducibility	0.00	0.00	0.00	0.00
Part to part variability	61.73	3810.82		72.04
Combined R&R	38.46	1478.91	100.00	27.96
Total	**72.73**	**5289.73**		**100.00**

Abbreviations: MS, Mean Score; R & R, Repeatability and Reproducibility; SD, Standard Deviation.

The largest uterus (patient 2) caused the most variation between estimated and real weight (Fig. **1**). The height of the boxes around the measurements shows the variation in each physician´s bias across the trials. The more experienced the physician was, the smaller was the variation. In addition, the more experienced the physician was in taking the ultrasound

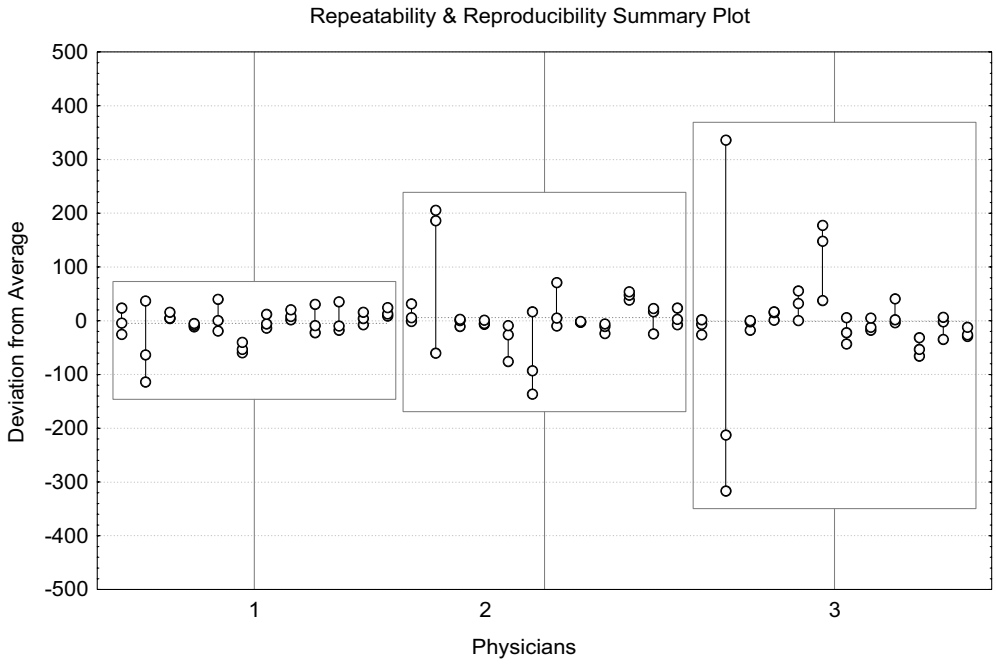

Fig. (1). Repeatability and reproducibility summary plot: difference of the real weight and estimated weight by combination formula. Patients are shown in study order by most experienced physician (1) to least experienced physician (3).

measurements the less variation there was between her own three measurements (Fig. **2**). The variation between real and estimated weight in patient 2 seems to be higher than the variation in other patients, which also showed some dependence on the experience of the physician (Fig. **3**).

If we remove the two patients with the largest uteruses (patients 2 and 6) from the study, the variation in the measurements of all the physicians varies very little compared to the original study setting. The variation between measurements (repeatability) falls to 22.4 %. Combined R&R falls to 23.7 %, also including the increase in variation of 1.4 % between physicians (reproducibility). Therefore the differences both between the three physicians and between their own measurements (average range 32 grams, warning limit 82 grams) fall. The differences in the physicians´ experience appears only in the measurements of the largest uteruses.

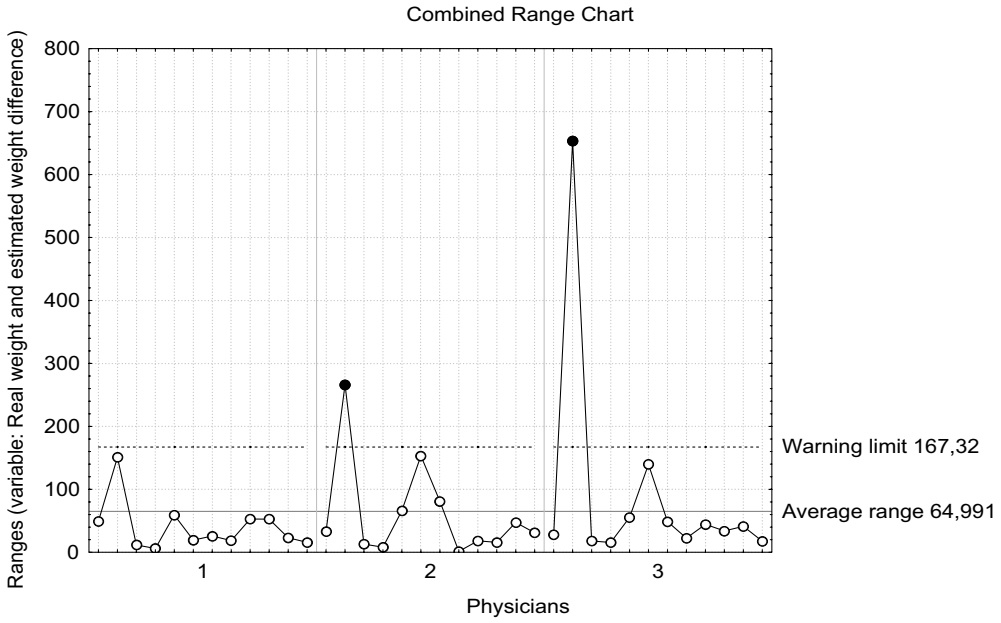

Fig. (2). Combined range chart: physicians by patients for difference of real weight and estimated weight by combination formula. Patients are shown in study order by most experienced physician (1) to least experienced physician (3). Warning limit presents three times standard deviation (99% of observations were below line).

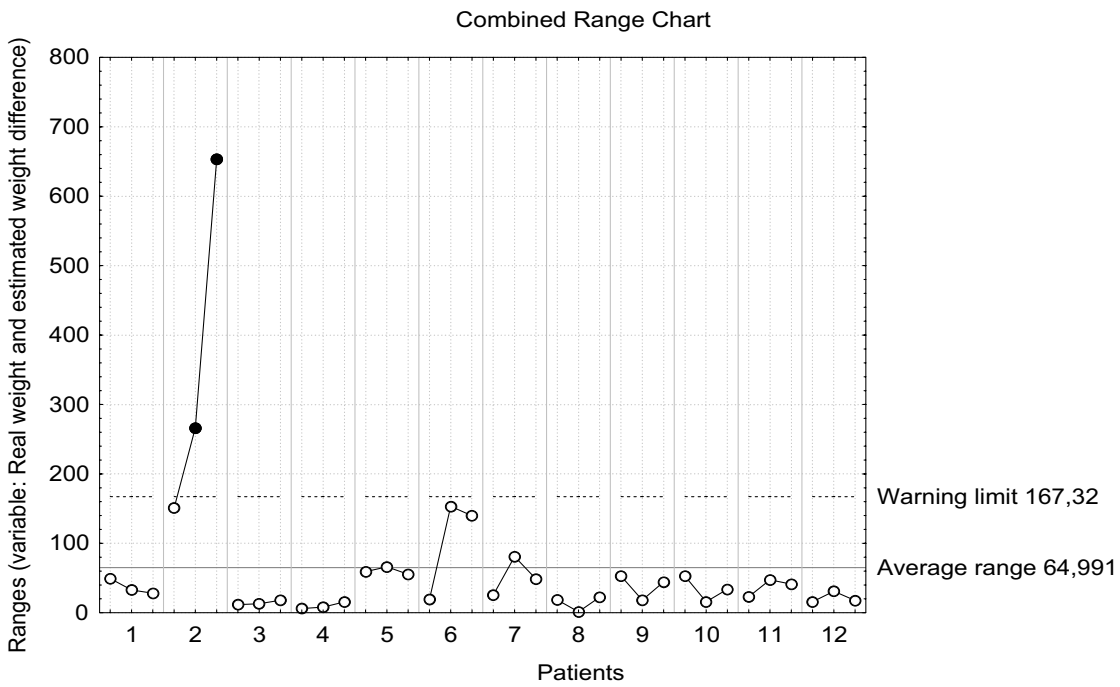

Fig. (3). Combined range chart: patients by physicians for difference of real weight and estimated weight by combination formula. Physicians are shown by most experienced to least experienced by patients shown in study order.

DISCUSSION

We studied patients with a myomatous uterus, as in clinical practice, they present most frequently for evaluation of uterine size, e.g. in relation to determination of the surgical approach to be used in hysterectomy [1]. We have recently reported that our formula combining the formulas for a prolate ellipsoid and a cylinder is more accurate in predicting the true total weight of the myomatous uterus than the traditional prolate ellipsoid formula alone in evaluating the dimensions of the uterine corpus and cervix with a transvaginal ultrasound probe [8].

We tested the quality of a new method for estimating the total volume of the uterus. The repeatability and reproducibility (R&R) method has rarely been used in medicine. Holli and her associates carried out a study on radiotherapy planning for breast cancer patients [10]. Eleven radiation oncologists planned radiotherapy three times for three different kinds of breast cancer patients without knowing that they were dealing with the same patient three times. Interphysician variation was not high but there were some clearly outlying physicians. The highest variation was in repeatability (intraphysician variation). Patient-to-patient variation accounted for the major part of the variation [10].

Merce and associates conducted a study to assess intraobserver and interobserver reproducibility of the parameters of ovarian response and oocyte ability and the influence of the ovarian functional stage [11]. They examined twenty-nine women with 3-dimensional ultrasonography and power Doppler angiography (PDA). They analysed ovarian volume, follicle number, vascularization index, flow index, and vascularization-flow index. They found excellent intraobserver and interobserver reproducibility for the ovarian volume, follicle counts, and 3-dimensional indices. The ovarian functional stage had no influence on reliability.

We utilized the repeatability and reproducibility method in training of ultrasound measurement of uterine volume. A combination of the volume of the uterine corpus and the volume of the cervix obtained by combining the formulas for a prolate ellipsoid and a cylinder appeared to be a useful means of estimating the actual volume of the total uterus. Despite the limitation of the small population size in this study the total number of findings (108) is sufficient for the R&R method.

Two sources of variation were identified: variation across trials (repeatability) and difference across physicians (reproducibility), and a combination of these. The more experienced the physician was in taking the ultrasound measurements the less variation there was between her own three measurements. Acceptable variation should be caused by patients. In the ideal case interphysician variation, intraphysician variation and combined repeatability and reproducibility should not exceed 10%. In our study variation between the measurements of each physician approached 30%. This may be an effect of the method of taking measurements by ultrasound. The ultrasound probe has to be rotated from the sagittal plane to the coronal axis of the organ during measurement. The probe may depart slightly from it´s original position during this rotation.

The two patients with the largest uteruses caused the most variation with estimated and real weight. The ultrasound scale of the vaginal probe was not wide enough to encompass the whole mass of the myomatous uterus in these patients and part of the measurements had to be taken with a transabdominal ultrasound probe. This may have affected the results. However, in the case of a large myomatous uterus it may be better to measure the uterine cervix with a transvaginal probe and the uterine corpus with a transabdominal probe. This method probably causes less variation between

estimated and real weight than use of the vaginal ultrasound probe alone.

CONCLUSIONS

We found that every investigator was able to perform a reliable measurement of the volume of the uterus by ultrasound. The experience of the physician had an effect on repeatability but not on reproducibility in estimating uterine weight by ultrasound. Experience diminished the variability between investigator´s own measurements. Our results indicate the importance of conscientiousness in taking measurements of the uterus and cervix.

ACKNOWLEDGEMENTS

This study was supported by grants from competitive research funding of the Pirkanmaa Hospital District, Finland.

REFERENCES

[1] Kovac SR. Guidelines to determine the route of hysterectomy. Obstet Gynecol 1995; 85: 18-23.

[2] Lepine LA, Hillis SD, Marchbanks PA, *et al*. Hysterectomy surveillance-United States, 1980-1993. MMWR 1997; 46: 1-15.

[3] Flickinger LD, D`Ablaing G, III, Mishell DR, Jr. Size and weight determinations of nongravid, enlarged uteri. Obstet Gynecol 1986; 68: 855-8.

[4] Saxton DW, Farquhar CM, Rae T, Beard RW, Anderson MC, Wadsworth J. Accuracy of ultrasound measurements of female pelvic organs. Br J Obstet Gynaecol 1990; 97: 695-9.

[5] Goldstein SR, Horii SC, Snyder JR, Raghavendra BN, Subramanyam B. Estimation of nongravid uterine volume based on nomogram of gravid uterine volume: Its value in gynecologic uterine abnormalities. Obstet Gynecol 1988; 72: 86-90.

[6] Kung F-T, Chang S-Y. The relationship between ultrasonic volume and actual weight of pathologic uterus. Gynecol Obstet Invest 1996; 42: 35-8.

[7] Becker R, Entezami M, Hese S, Vollert W, Loy V, Weitzel HK. Gynäkologische Sonographie II. Normalwerte der standardisierten Uterusbiometrie. Ultraschall Klin Prax 1994; 8: 241-7.

[8] Rovio PH, Luukkaala T, Vuento M, Oksa S, Sundström H, Heinonen PK. Ultrasonographic weight of the myomatous uterus: a pilot study assessing a new combined geometrical formula. Eur J Obstet Gynecol Reprod Biol 2008; 137: 193-7.

[9] Duncan AJ. Quality control and industrial statistics. 5th ed. Homewood, IL:Irwin; 1986.

[10] Holli K, Laippala P, Ojala A, Pitkänen M. Quality control in health care: an experiment in radiotherapy planning for breast cancer patients after mastectomy. Int J Radiat Oncol Biol Phys 1999; 44: 827-33.

[11] Merce LT, Gomez B, Engels V, Bau S, Bajo JM. Intraobserver and interobserver reproducibility of ovarian volume, antral follicle count, and vascularity indices obtained with transvaginal 3-dimensional ultrasonography, power Doppler angiography, and the virtual organ computer-aided analysis imaging program. J Ultrasound Med 2005; 24: 1279-87.

Received: March 5, 2008 Revised: June 4, 2008 Accepted: June 11, 2008