DR ANNE GISSELMANN EGEKVIST (Orcid ID : 0000-0001-8346-932X)

PROFESSOR ULRIK SCHIØLER KESMODEL (Orcid ID : 0000-0003-3868-106X)

Intra- and interobserver variability in nodule size of rectosigmoid endometriosis measured by two- and three-dimensional transvaginal sonography

Running headline: Rectosigmoid endometriosis nodule variability

Anne Gisselmann EGEKVIST[1,2], Axel FORMAN[1,2], Mads RIISKJÆR[1,2], Ulrik Schiøler KESMODEL[3,4], Mie MATHIASEN[1,2] & Mikkel SEYER-HANSEN[1,2]

[1]Department of Clinical Medicine, Aarhus University, Aarhus,[2]Department of Obstetrics and Gynecology, Aarhus University Hospital, Aarhus, [3]Department of Obstetrics and Gynecology, Herlev and Gentofte University Hospital, Herlev, [4] Institute for Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:

Anne Gisselmann Egekvist

Aarhus University, Department of Clinical Medicine/Department of Obstetrics and Gynecology, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark

E-mail: ageg@clin.au.dk

Conflicts of Interest notification

None of the authors have any conflicts of interest.

Abstract

Introduction: The aim of the study was to assess the intra- and interobserver variability of two- and three-dimensional rectosigmoid nodule size measurements by transvaginal ultrasonography in patients with rectosigmoid endometriosis. *Material and methods:* Intra- and interobserver variability was assessed in 10 and 30 patients, respectively. Measurements in two dimensions were performed in real-time during the scan, and three-dimensional measurements of volume were done on a computer. Differences within and between observers were expressed in absolute units (mm) and percentage (%) of average nodule size. Coefficient of repeatability and Bland–Altman plots with limits of agreement were used to evaluate the intra- and interobserver variability. *Results:* Intra- and interobserver variability in two-dimensional sonography ranged from 11-14 mm or 46-51 % for length, 3-6 mm or 32-57 % for depth and 5-9 mm or 33-58 % for width of the nodule. Results of three-dimensional sonography, with assessment of nodule volume, showed intra- and interobserver variability between 0.4-2.5 times the average nodule size. *Conclusions:* Measurements of rectosigmoid endometriosis nodule size with two- and three- dimensional transvaginal ultrasonography were associated with large intra- and interobserver variability, and these techniques should therefore be used with caution in clinical control and research of nodule growth.

Keywords

Rectosigmoid endometriosis, Ultrasonography, Transvaginal, Observer variation, deeply infiltrating endometriosis, endometriosis

Abbreviations

2D: two-dimensional
3D: three-dimensional
AGE: Anne Gisselmann Egekvist
CI: confidence interval
CoR: coefficient of repeatability

DIE: deeply infiltrating endometriosis

LoA: limits of agreement

MR: Mads Riiskjær

MSH: Mikkel Seyer-Hansen

SD: standard deviation

TVS: transvaginal sonography

VOCAL: Virtual Organ Computer-aided Analysis

Key Message

The intra- and interobserver variabilities when measuring rectosigmoid endometriosis with transvaginal ultrasonography in two and three dimensions, was evaluated. Large variabilities were found on all parameters. This has to be taken into account when evaluating size of the lesion.

Introduction

Endometriosis is characterized by the occurrence of endometrial-like tissue outside the uterus and represents a common cause of chronic pelvic pain (1). Deeply infiltrating endometriosis (DIE) denotes endometriosis invading more than 5 mm under the peritoneum (2). A common site for DIE is the rectosigmoid bowel, and patients with this phenotype often suffer from severe dysmenorrhea, dyschezia and dyspareunia (3).

Management of patients with rectosigmoid DIE is widely discussed (3-6), since surgery is associated with risk of major complications (7). This mainly concerns segmental resection (8), while local excision limited to the ventral wall of the bowel may be associated with reduced risk (9). Conservative treatment with oral contraceptives, the levonorgestrel intrauterine device or oral gestagens is therefore the primary approach in many centers, while discussion of possibilities for surgery is reserved for patients with severe side effects, poor symptom control or wish for pregnancy. Detailed information on the risk of worsened disease at a later stage is pertinent in these cases, but little is known about the potential growth of rectosigmoid DIE during conservative treatment. These patients are therefore offered annual control with clinical examination and perhaps transvaginal sonography (TVS) (10, 11).

Two-dimensional (2D) TVS has high sensitivity and specificity in the diagnosis of rectosigmoid endometriosis (12) and the method is recommended as a first-line diagnostic tool (13, 14) . Three-dimensional (3D) TVS has been introduced as an alternative (15), and both methods have been used before and after treatment to evaluate efficacy of medical therapy (16, 17). Hence, 2D and/or 3D TVS could represent an ideal objective parameter, but the approach requires that the intra- and interobserver variabilities for measurements of nodule dimension are known.

Our group previously reported on the interobserver variability in 2D between an experienced and less experienced observer for the size of rectosigmoid DIE measured with TVS (18). In this study we examined the intra- and interobserver variability for 2D and 3D TVS measurements of rectosigmoid nodule dimensions and volume.

Material and methods

Between October 2014 and August 2016 patients were enrolled for the study at the Department of Obstetrics and Gynecology, Aarhus University Hospital, which is one of two tertiary referral centers for endometriosis in Denmark.

Inclusion criteria were the presence of rectosigmoid DIE confirmed by TVS or magnetic resonance imaging. Patients were included when participating observers (Anne Gisselmann Egekvist (AGE), Mads Riiskjær (MR), Mikkel Seyer-Hansen (MSH)) were present at the hospital. Exclusion criteria were patients where TVS could not be performed due to pain, cases where the oral border of the nodule could not be seen or if the nodule was wedged between endometriomas.

Three members of the subspecialized endometriosis team performed the examinations. MSH had 14 years of experience with TVS in gynecology and previously cooperated on interobserver variability in TVS assessment of dimensions of nodules with AGE, who had five years of experience (18). MR had seven years of experience with TVS in gynecology, whereof two years had been with rectosigmoid endometriosis.

The study was based on pairwise measurements in 2D and 3D between the three observers (Figure 1).

Examinations were performed with the same Voluson® E8 (GE Healthcare, Zipf, Austria) with a 6-12 MHz vaginal probe. No bowel preparation was used. An acoustic window was created with a medium sized glove as probe cover, with gel in the tip of the middle finger. At

first the rectal wall was identified as a thin hypoechoic line in close relation to the posterior vaginal wall. The bowel wall was followed cranially until the rectosigmoid nodule presented as a regular or irregular hypoechoic mass >3 mm, replacing the regular appearance of the muscular layer (Figure 2) (11, 19, 20). The upper limit for detection of bowel nodules was set to the rectosigmoid junction. Therefore, nodules examined in this study are located below this level. In cases where more than one nodule was present, the lesion closest to the anal verge was measured.

2D TVS

We measured length, depth and width of the rectosigmoid nodule (Figure 2). Depth of infiltration was measured antero-posterior to the plane where maximal length along the longitudinal axis of the rectosigmoid bowel could be displayed. When depth of infiltration had been marked on the image, length was measured by adding the distances from this line to the anal and oral peripheral limits of the nodule, in order to compensate for bowel angulation. For length, the peripheral limits were defined by thickness of the muscular layer ≥3 mm (Figure 2a). For measurement of width, delimitations of the nodule were defined by the lateral limits of the infiltrated muscular layer in an image transverse to the bowel axis (Figure 2b). All measurements were made in three orthogonal planes (14). Only the part of the nodule located in the muscular layer was measured since TVS is connected to low sensitivity evaluating infiltration of the submucosal and mucosal layer (19). Measurements of rectosigmoid endometriosis in 2D were made in real-time, during the examination. All images were stored on the ultrasound machine and saved to a computer.

3D TVS

After measurement of the rectosigmoid nodule in two dimensions each of the observers obtained a 3D dataset for assessment of intra- and inter-observer variability in three dimensions. The nodule was centered longitudinally in the image. Thereafter, a 3D dataset was obtained using a sweep angel of 120° and maximal quality setting. The datasets were transferred to a computer and volume measurements were done in the 4D View® program (GE Healthcare, Milwaukee, WI).

Calculations were performed by Virtual Organ Computer-aided Analysis (VOCAL®). In the VOCAL® setting the dataset can be rotated around a fixed axis and in each rotation-step the region of interest can be outlined. The program was set to rotate the dataset 9° around the B-plane (vertical axis) with volume calculations based on 20 planes (21).

Intraobserver variability

The patients underwent two TVS examinations (performed by AGE) with an interval of less than 14 days irrespective of the cycle day. Results from the first examination were not available at the second procedure.

Interobserver variability

In the assessment of interobserver variability observers worked in two pairs on separate patient groups: MR and AGE (MR:AGE) or MSH and AGE (MSH:AGE).

In each of the two pairs, the second observer was unaware of the measurements of the first observer by blinding measurements and vision to the screen. For measurements of volume in the 3D dataset observers were not present in the same room.

Statistical analyses

We calculated the mean differences between measurements, 95% confidence intervals (CI) and the standard deviation (SD) (22). Histograms and quantile-quantile plots were made to test for normality. If the data were not normally distributed, analysis on log scale was performed. Measurements were marked in Bland-Altman plots showing differences against the average of the measurements for each patient. The SD of the mean difference between observers was used to calculate 95% limits of agreement (LoA) (mean +/- 1.96 x SD). This is an interval where 95% of all the differences should be contained if the data are normally distributed (23). The mean difference between measurements of the observers is an estimate of the average bias and Student paired t-test was used to test systematic bias between observers. It is a prerequisite for evaluation of inter- and intraobserver variability that the measurements do not differ systematically in which case the dataset cannot not be used to estimate measurement error [22].

We used the Coefficient of Repeatability (CoR) expressed as 1.96 x SD (23). CoR represents the difference between two measurements that will be exceeded by only 5% of the differences between pairs of measurements (22).

No power-calculations were performed. A power calculation requires a meaningful statistical hypothesis that can be tested, and as the aim was simply to observe and describe whatever differences exist, no one specific hypothesis would seem meaningful. The number of

participants to be included in the study was therefore based on what was possible in our clinic and on looking through other intra- and interobserver studies.

Ethical approval

All patients gave informed consent and the study was approved by The Central Denmark Region Committees on Health Research Ethics (no. 1-10-72-196-13) and the Danish Data Protection Agency (no. 1-16-2-657-15).

Results

Ten patients were included in the intraobserver variability study. For each pair of observers in the interobserver variability study, 30 patients were included.

Clinical data and average dimensions of nodules are presented in table 1. Patients included by MR:AGE had a tendency towards larger body mass index and larger nodules compared to patients included by AGE:AGE and MSH:AGE.

2D TVS

All data in the 2D intra- and interobserver analysis were normally distributed and differences between measurements are given in mm. Length, depth and width were correlated (data not shown).

The mean difference between measurements, 95% CI, CoR, LoA and p-values for test of systematic difference regarding length, depth and width, in absolute and relative measures, are shown in table 2.

No systematic differences were found between measurements in any of the pairs of observers. Bland-Altman plots with mean and LoA for the intra-and interobserver variability of length, depth and width are presented in Figure 3 and 4. This plot depicts the differences between measurements in mm against the average of the measurements in mm for each patient. There was no tendency towards increasing SD with increasing size of the nodule (except width for MR:AGE) i.e. variability did not vary with object size. Therefore, the LoA and CoR were reasonable estimates of the variability between all measurements, independent of the value of the single measurement.

For length, the intraobserver CoR (AGE:AGE) was 11.0 mm, meaning that 5% of differences in length measurement exceeded this value (Table 2). The interobserver CoR amounted to

13.3 mm and 14.1 mm for MR:AGE and MSH:AGE, respectively. Thus, CoR in percent of the average length ranged from 46 to 51 %.

For infiltration depth, intraobserver CoR (AGE:AGE) was 2.9 mm. Interobserver variabilities were 5.5 mm (MR:AGE) and 2.7 mm (MSH:AE). Thus, the interobserver CoR was larger for MR:AGE compared to AGE:AGE and MSH:AGE. CoR in percent of the average depth ranged from 32 to 57%.

For width of the nodule, intraobserver (AGE:AGE) CoR was 4.8 mm. Interobserver variabilities were CoR 9.2 mm (MR:AGE) and 5.4 mm (MSH:AGE). CoR in percent of the average width ranged from 33 to 58%.

In the intraobserver variability study, all but two patients were treated medically, three had levonorgestrel intrauterine device, three had oral contraceptives, one had oral gestagens and one had combined oral gestagens and levonorgestrel intrauterine device. Calculations were performed with and without inclusion of the two untreated patients, and this did not change the results.

3D TVS

For the 3D intra- and interobserver analysis of volume, data did not follow a normal distribution and were logtransformed. Back-transformation was performed before creating the Bland-Altman plots and the resulting y-axis depicts proportions of differences. The mean difference and LoA should therefore be interpreted as relative measures, and CoR could not be used as an absolute estimate of the measurement error. In Figure 5, Bland-Altman plots with mean and LoA for the intra-and interobserver variability of volume are presented. No tendency towards increasing SD with increasing size of the nodule was found.

For MR:AGE, one observer systematically made larger measurements than the other (Table 3). No significant difference was found in AGE:AGE and MSH:AGE.

LoA for AGE:AGE and MSH:AGE were almost identical. Interpretation of these data, achieved after backtransformation from logtransformed data, is that a volume measured by the first observer would be between 0.4 times and 2.4 times the volume found by the second observer for 95% of the observations.

Discussion

The present study was designed to determine the intra- and inter-observer variability in 2D and 3D measurements of rectosigmoid DIE dimensions and volume by TVS. These methods differ in methodology and will be discussed separately.

CoR reflects the variability within and between observers and only 5 % of measurement differences will exceed this value (22). This simplifies quantification of intra- and interobserver variability, which is determined by multiple factors such as image quality, ultrasound resolution, nodule irregularity, difficulty level of the TVS and experience of the investigator.

High CoR values were found for length, depth and width of the nodules. This was true for both intra- and interobserver variabilities. In order to achieve a rough estimation of the clinical usefulness of TVS in assessment of DIE growth, CoR for each dimension was expressed in percent of the average measurements. This revealed that CoR was at least one third of nodule size for length, depth and width.

Length of the bowel nodule is important for decisions on surgical technique. An upper limit of 20-30 mm has been suggested when the circular stapler is used (24, 25). However, if the distance from the anal verge is below 10 cm the Rouen method could represent an alternative where local resections can be performed for a large range of up to 50 mm in length (26). For such dimensions, our results might suggest less relative uncertainty compared to small-sized nodules but studies with larger numbers of nodules with length >30 mm are needed to address this issue.

Depth and width of infiltration relates to the grade of bowel stenosis (27). In general, the opportunity for disc resection declines if the stenosis grade exceed 50% [24]. The use of TVS in this context is uncertain with the present variabilities for measurement of depth and width of nodules.

Factors behind the high CoR values for length and width may relate to ill-defined histopathological margins of DIE with offshoots into surrounding tissues (28). This makes exact delineation difficult despite our attempt to use objective definitions.

The findings in this study indicate that with current TVS techniques, length, depth and width measurements are associated with large intra- and interobserver variability.

It has been well documented that TVS represents a method with high positive and negative predictive values in the qualitative diagnosis of rectosigmoid DIE (29, 30). However, our

findings regarding size of the lesion leaves the clinician with uncertainty when growth is suspected at repeated clinical control. Magnetic resonance might be suggested, but no studies has yet been performed on the intra- and interobserver variability for measurements of DIE nodule size with this method and the value of this approach for detection of DIE growth need further investigation.

We found large intra- and interobserver variabilities when measuring volume of rectosigmoid endometriosis by use of VOCAL®. The intra- and inter-observer variability was identical. An explanation may be that use of the VOCAL-technique implies integration of 20 areas that are manually delineated and may be difficult to define (11). This could introduce significant random errors.

In a previous study [17], the coefficient of variation for volume assessment of rectovaginal DIE infiltrating the rectum amounted to 0.6-2.2 %, based on three analyses of 10 recordings. Our results obtained with the same ultrasound device could not confirm these findings. Some of the discrepancy might be caused by differences in the volumes recorded, which were larger (3.5 cm^3 (17)) compared to our study (1.7 cm^3), and the use of bowel preparation. Our study restricted volume recording to the muscular layer of the bowel, as opposed to the total nodule (17). Still, however, our results motivate caution in the interpretation of changes in DIE volume when measured with the VOCAL-technique.

In this study, observers measured dimensions of rectosigmoid DIE in 2D-realtime (once on each nodule). Our setting reflects clinical practice where one single measurement is usually made.

It might be argued that CoR in percent of average dimension would be less for large nodules as supported by the Bland-Altman plots (Figure 3 and 4). The distribution of nodule size was not wide enough to allow for such analysis, and further studies including patients with larger nodules are needed to evaluate this aspect.

Systematic differences in the assessments were found especially for MR:AGE, which may be due to the fact that this pair of observers had not previously worked together and MR not being as experienced with the TVS technique as MSH and AGE. Thus, this most likely reflects the clinical everyday setting.

This study illustrated the intra- and interobserver variability of one observer and two pairs of observers, respectively. A larger number of observers would be preferable, though this study

was conducted in an outpatient clinic with real-time TVS and the study design had to be practicable.

Conclusions

Our results showed large intra- and interobserver variabilities when 2D or 3D TVS are used for measurements of rectosigmoid DIE dimensions. Although, these techniques are pertinent to the qualitative diagnosis of the disease, they should be used with caution when changes in dimension are to be assessed in clinical follow-up and research on rectosigmoid DIE growth.

Funding

References

1.   Giudice LC. Clinical practice. Endometriosis. N Engl J Med. 2010;362:2389-98.

2.   Cornillie FJ, Oosterlynck D, Lauweryns JM, Koninckx PR. Deeply infiltrating pelvic endometriosis: histology and clinical significance. Fertil Steril. 1990;53:978-83.

3.   Vercellini P, Crosignani PG, Somigliana E, Berlanda N, Barbara G, Fedele L. Medical treatment for rectovaginal endometriosis: what is the evidence? Hum Reprod. 2009;24:2504-14.

4.   Bassi MA, Podgaec S, Dias JA, Jr., D'Amico Filho N, Petta CA, Abrao MS. Quality of life after segmental resection of the rectosigmoid by laparoscopy in patients with deep infiltrating endometriosis with bowel involvement. J Minim Invasive Gynecol. 2011;18:730-3.

5.   Vercellini P, Crosignani PG, Abbiati A, Somigliana E, Vigano P, Fedele L. The effect of surgery for symptomatic endometriosis: the other side of the story. Hum Reprod Update. 2009;15:177-88.

6. Koninckx PR, Ussia A, Adamyan L, Wattiez A, Donnez J. Deep endometriosis: definition, diagnosis, and treatment. Fertil Steril. 2012;98:564-71.

7. Riiskjaer M, Forman A, Kesmodel US, Andersen LM, Ljungmann K, Seyer-Hansen M. Diagnostic Value of Serial Measurement of C-Reactive Protein in the Detection of a Surgical Complication after Laparoscopic Bowel Resection for Endometriosis. Gynecol Obstet Invest. 2017;82:410-6.

8. Riiskjaer M, Greisen S, Glavind-Kristensen M, Kesmodel US, Forman A, Seyer-Hansen M. Pelvic organ function before and after laparoscopic bowel resection for rectosigmoid endometriosis: a prospective, observational study. BJOG. 2016;123:1360-7.

9. Woods RJ, Heriot AG, Chen FC. Anterior rectal wall excision for endometriosis using the circular stapler. ANZ J Surg. 2003;73):647-8.

10. Abrao MS, Petraglia F, Falcone T, Keckstein J, Osuga Y, Chapron C. Deep endometriosis infiltrating the recto-sigmoid: critical factors to consider before management. Hum Reprod Update. 2015;21:329-39.

11. Koga K, Osuga Y, Yano T, Momoeda M, Yoshino O, Hirota Y, et al. Characteristic images of deeply infiltrating rectosigmoid endometriosis on transvaginal and transrectal ultrasonography. Hum Reprod .2003;18:1328-33.

12. Hudelist G, English J, Thomas AE, Tinelli A, Singer CF, Keckstein J. Diagnostic accuracy of transvaginal ultrasound for non-invasive diagnosis of bowel endometriosis: systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2011;37:257-63.

13. Piketty M, Chopin N, Dousset B, Millischer-Bellaische AE, Roseau G, Leconte M, et al. Preoperative work-up for patients with deeply infiltrating endometriosis: transvaginal ultrasonography must definitely be the first-line imaging examination. HumReprod. 2009;24:602-7.

14. Guerriero S, Condous G, Van den Bosch T, Valentin L, Leone FP, Van Schoubroeck D, et al. Systematic approach to sonographic evaluation of the pelvis in women with suspected endometriosis, including terms, definitions and measurements: a consensus opinion from the International Deep Endometriosis Analysis (IDEA) group. Ultrasound Obstet Gynecol. 2016;48:318-32.

15. Guerriero S, Saba L, Ajossa S, Peddes C, Angiolucci M, Perniciano M, et al. Three-dimensional ultrasonography in the diagnosis of deep endometriosis. Hum Reprod. 2014;29:1189-98.

16. Mabrouk M, Frasca C, Geraci E, Montanari G, Ferrini G, Raimondo D, et al. Combined oral contraceptive therapy in women with posterior deep infiltrating endometriosis. J Minim Invasive Gynecol. 2011;18:470-4.

17. Ferrero S, Leone Roberti Maggiore U, Scala C, Di Luca M, Venturini PL, Remorgida V. Changes in the size of rectovaginal endometriotic nodules infiltrating the rectum during hormonal therapies. Arch Gynecol Obstet. 2013;287:447-53.

18. Egekvist AG, Forman A, Seyer-Hansen M. Transvaginal ultrasonography of rectosigmoid endometriosis: interobserver variation of lesion size. Acta Obstet Gynecol Scand. 2012;91:264-8.

19. Hudelist G, Tuttlies F, Rauter G, Pucher S, Keckstein J. Can transvaginal sonography predict infiltration depth in patients with deep infiltrating endometriosis of the rectum? Hum Reprod. 2009;24:1012-7.

20. Bazot M, Malzy P, Cortez A, Roseau G, Amouyal P, Darai E. Accuracy of transvaginal sonography and rectal endoscopic sonography in the diagnosis of deep infiltrating endometriosis. Ultrasound ObstetGynecol. 2007;30:994-1001.

21. Raine-Fenning N, Campbell B, Collier J, Brincat M, Johnson I. The reproducibility of endometrial volume acquisition and measurement with the VOCAL-imaging program. Ultrasound ObstetGynecol. 2002;19:69-75.

22. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol. 2003;22:85-93.

23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307-10.

24. Landi S, Pontrelli G, Surico D, Ruffo G, Benini M, Soriano D, et al. Laparoscopic disk resection for bowel endometriosis using a circular stapler and a new endoscopic method to control postoperative bleeding from the stapler line. J Am Coll Surg. 2008;207:205-9.

25. de Almeida A, Fernandes LF, Averbach M, Abrao MS. Disc resection is the first option in the management of rectal endometriosis for unifocal lesions with less than 3 centimeters of longitudinal diameter. Surg Technol Int. 2014;24:243-8.

26. Roman H, Abo C, Huet E, Bridoux V, Auber M, Oden S, et al. Full-Thickness Disc Excision in Deep Endometriotic Nodules of the Rectum: A Prospective Cohort. Dis Colon Rectum. 2015;58:957-66.

27. Abrao MS, Podgaec S, Dias JA, Jr., Averbach M, Silva LF, Marino de Carvalho F. Endometriosis lesions that compromise the rectum deeper than the inner muscularis

layer have more than 40% of the circumference of the rectum affected by the disease. J Minim Invasive Gynecol. 2008;15:280-5.

28. Anaf V, El Nakadi I, Simon P, Van de Stadt J, Fayt I, Simonart T, et al. Preferential infiltration of large bowel endometriosis along the nerves of the colon. Hum Reprod. 2004;19:996-1002.

29. Guerriero S, Ajossa S, Gerada M, Virgilio B, Angioni S, Melis GB. Diagnostic value of transvaginal 'tenderness-guided' ultrasonography for the prediction of location of deep endometriosis. Hum Reprod. 2008;23:2452-7.

30. Abrao MS, Goncalves MO, Dias JA, Jr., Podgaec S, Chamie LP, Blasbalg R. Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. Hum Reprod. 2007;22:3092-7.

Legends for figures

Figure 1. Overview of recorded images and assessment of intra- and interobserver variability. AGE: Anne Gisselmann Egekvist; MR:Mads Riiskjær; MSH: Mikkel Seyer-Hansen.

Figure 2. Placement of calibers for rectosigmoid nodule measurements. (a) 1 and 2: Peripheral limits for length defined by thickness of the muscular layer ≥3 mm. 3: Depth of infiltration. 4+5: Length of the nodule accounting for bowel angulation. (b) 1: Width of the nodule.

Figure 3. Intraobserver variability in 2D length (a), depth (b) and width (c) measurements of rectosigmoid endometriosis for Anne Gisselmann Egekvist (AGE). Short-dashed line: Mean difference, long-dashed line: Upper- and lower limit of agreement, solid line: Zero.

Figure 4. Interobserver variability in 2D length (a), depth (b) and width (c) measurements for Mads Riiskjær (MR) and Anne Gisselmann Egekvist (AGE) (MR:AGE) and length (d), depth (e) and width (f) measurements for Mikkel Seyer-Hansen (MSH) and Anne Gisselmann Egekvist (AGE) (MSH:AGE) of rectosigmoid endometriosis. Short-dashed line: Mean difference, long-dashed line: Upper- and lower limit of agreement, solid line: Zero

Figure 5. Intraobserver variability for Anne Gisselmann Egekvist (AGE) (AGE:AGE) (a). Interobserver variability for Mads Riiskjær (MR) and Anne Gisselmann Egekvist (AGE) MR:AGE (b). Interobserver variability for Mikkel Seyer-Hansen (MSH) and Anne Gisselmann Egekvist (AGE) MSH:AGE (c) of 3D volume measurements of rectosigmoid endometriosis. Short-dashed line: Mean difference, long-dashed line: Upper- and lower limit of agreement, solid line: One

Table 1 Demographic data and average size of rectosigmoid nodules in 2D and 3D. Numbers are given in mean (SD) unless otherwise specified.

2D	AGE:AGE (n=10)	MR:AGE (n=30)	MSH:AGE (n=30)
Age	38.2 (6.2)	37.5 (6.2)	37.2 (7.6)
Body mass index	23.5 (3.5)	26.8 (4.8)*	24.1 (4.4)*
Average nodule dimension (mm)			
Length	26.6 (7.5)	31.3 (10.3)	29.1 (11.6)
Depth	8 (2.0)	10.2 (2.6)	9.3 (3.4)
Width	14 (4.7)	14.9 (3.5)	13.6 (4.5)

	3D		
Average nodule volume (ml)			
Median (10th-90th percentile)	1.7 (1.0 – 7.0)	2.9 (1.0 – 6.9)	2.1 (0.6 – 8.6)

AGE:AGE: Intraobserver variability for Anne Gisselmann Egekvist, MR:AGE: Interobserver variability for Mads Riiskjær and Anne Gisselmann Egekvist, MSH:AGE: Interobserver variability for Mikkel Seyer-Hansen and Anne Gisselmann Egekvist, *Height and weight missing on five patients in MR:AGE and three patients in MSH:AGE. Values imputed from patient characteristics in another study.

Table 2 Intra- and interobserver variability measuring rectosigmoid nodule dimension in 2D.

2D		Absolute difference (mm) between measurements				Relative (%) difference between measurements		
	Parameter	Mean (95 % CI)	CoR (1.96 x SD)	LoA	p-value	Mean (95 % CI)	CoR (1.96 x SD)	LoA
Intraobserver variability								
AGE:AGE (n=10)								
	Length	-2.1 (-6.1 - 1.9)	11.0	-13.1 - 9.0	0.270	-7.6 (-24.5 – 9.2)	46.1	-53.8 - 38.5
	Depth	0.6 (-0.4 –1.6)	2.9	-2.4 - 3.6	0.239	7.6 (-5.9 – 21.1)	37.0	-29.4 - 44.6
	Width	1.6 (-0.1 – 3.3)	4.8	-3.1 - 6.3	0.065	11.5 (-0.7 – 23.7)	33.4	-22.0 - 45.0

Interobserver variability							
MR:AGE (n=30)							
Length	0.6 (-2.0 – 3.1)	13.3	-12.8-14.0	0.653	1.8 (-7 – 11.6)	51.4	-49.5 - 53.2
Depth	0.9 (-0.2 – 1.9)	5.5	-4.5 - 6.3	0.100	8.3 (-2.6 – 19.2)	57.1	-48.7 - 65.4
Width	0 (-1.7 – 1.7)	9.2	-9.1 - 9.1	1.000	0 (-11.5 10.6)	58.0	-58.4 - 57.5
MSH:AGE (n=30)							
Length	-1.0 (-3.7 – 1.7)	14.1	-15.1-13.1	0.438	-1.2 (-10.6 – 8.2)	49.4	-50.6 - 48.3
Depth	0.4 (-0.2 – 0.9)	2.7	-2.4 - 3.2	0.170	5.1 (-1.0 – 11.2)	32.2	-27.1 - 37.3

Width	-0.2	5.4	-5.5 - 5.2	0.739	-0.4	40.8	-41.2 - 40.5
	(-1.2 – 0.8)				(-8.2 – 7.4)		

CI: Confidence Interval, CoR: Coefficient of Repeatability, LoA: Limits of Agreement, AGE:AGE: Intraobserver variability for Anne Gisselmann Egekvist, MR:AGE: Interobserver variability for Mads Riiskjær and Anne Gisselmann Egekvist, MSH:AGE: Interobserver variability for Mikkel Seyer-Hansen and Anne Gisselmann Egekvist.

Table 3 Intra- and Interobserver variability when measuring rectosigmoid nodule volume in 3D. Data were logtransformed, and values should be interpreted as proportions.

3D	Relative difference (proportion) between measurements of volume (ml)		
	Mean (95 % CI)	LoA	p-value
Intraobserver variability			
AGE:AGE (n=10)	1.0 (0.7 – 1.4)	0.4 – 2.5	0.945
Interobserver variability			
MR:AGE (n=30)	1.2 (1.0 – 1.6)	0.4 – 4.0	0.022
MSH:AGE (n=30)	1.0 (0.8 – 1.2)	0.4 – 2.4	0.849

CI: Confidence Interval, LoA: Limits of Agreement, AGE:AGE: Intraobserver variability for Anne Gisselmann Egekvist, MR:AGE: Interobserver variability for Mads Riiskjær and Anne Gisselmann Egekvist, MSH:AGE: Interobserver variability for Mikkel Seyer-Hansen and Anne Gisselmann Egekvist

2D and 3D intra- and interobserver variability

(a) n=10

TVS AGE

Time: ≤ 14 days

TVS AGE

2D image
Length, depth and width (mm)

3D image
Volume (ml)

2D image
Length, depth and width (mm)

3D image
Volume (ml)

2D intraobserver variability
AGE:AGE

3D intraobserver variability
AGE:AGE

(b) n=30

TVS AGE

TVS MR

2D image
Length, depth and width (mm)

3D image
Volume (ml)

2D image
Length, depth and width (mm)

3D image
Volume (ml)

2D interobserver variability
MR:AGE

3D interobserver variability
MR:AGE

(c) n=30

TVS AGE

TVS MSH

2D image
Length, depth and width (mm)

3D image
Volume (ml)

2D image
Length, depth and width (mm)

3D image
Volume (ml)

2D interobserver variability
MSH:AGE

3D interobserver variability
MSH:AGE

(a)

(b)

(c)

(a)

(b)

(c)