ISUOG

Virtual World Congress

ON ULTRASOUND IN OBSTETRICS AND GYNECOLOGY

16–18 OCTOBER 2020

3 DAY EVENT

5 streams — **170+** expert talks — **1200+** abstracts

Register to access a high tech virtual space enabling you to learn and interact with

200+ world leading experts

and global community

Watch Live OR **On Demand**

Earn CME/CPD points for attending

Benefit from reduced registration fees

Explore scientific program & register here

CELEBRATING THIRTY YEARS • ISUOG WORLD CONGRESS • 30

Ultrasound diagnosis of endometrial cancer by subjective pattern recognition in women with postmenopausal bleeding: a prospective inter-rater agreement and reliability study

M. Wong[1], N. Thanatsis[1], T. Amin[1], E. Bean[1], K. Madhvani[2], D. Jurkovic[1]

[1]Institute for Women's Health, University College London Hospitals, London, UK

[2]Department of Gynaecology, Poole Hospital, Dorset, UK

Correspondence:

Mr Davor Jurkovic

Department of Obstetrics and Gynaecology

University College London Hospitals

235 Euston Road, London, NW1 2BU

Email: davor.jurkovic@nhs.net

Short title: Reliability of diagnosing endometrial cancer

Keywords: postmenopausal bleeding, endometrial cancer, ultrasound pattern recognition, ultrasound subjective assessment, inter-rater agreement, inter-rater reliability

Contribution

What are the novel findings of this work?

This is the first reported study where inter-rater reliability of diagnosing endometrial cancer on ultrasound by subjective pattern recognition is tested with individual acquisitions. Our results

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.22141

showed that there was good inter-rater reliability in diagnosing endometrial cancer on ultrasound in women with postmenopausal bleeding.

What are the clinical implications of this work?

Our findings should facilitate the wider use of subjective pattern recognition for endometrial cancer in routine clinical practice. Future research should evaluate the efficacy of subjective pattern recognition in prioritising women with likely endometrial cancer for urgent histological confirmation and referral to a gynaecological oncology centre.

Abstract

Objectives

The objective of our study was to assess the inter-rater agreement and reliability of using subjective pattern recognition to diagnose endometrial cancer on ultrasound in women with postmenopausal bleeding.

Methods

This was a prospective cross-sectional study at a gynaecological rapid access clinic, between Oct 2016 – Dec 2017, where consecutive women with postmenopausal bleeding and endometrial thickness of ≥4.5mm on transvaginal ultrasound scan were included. We excluded women on hormone replacement therapy, tamoxifen or with a history of primary gynaecological malignancy. Two raters independently performed ultrasound examinations, blinded to each other's findings, and classified women into having uniformly thickened endometrium, benign polyp or endometrial cancer, by subjective pattern recognition. Inter-rater reliability of the ultrasound diagnoses was assessed by using Cohen's kappa statistics. All women subsequently underwent either outpatient endometrial biopsy, hysteroscopy or hysterectomy.

Results

Forty women were included in the study with a median age of 61 (IQR 57-69) and a median endometrial thickness of 11.0mm (IQR 6.2-20.3). Final histological analysis confirmed 16 (40%) women with endometrial cancer, 16 (40%) endometrial polyps, 4 (10%) atrophic endometrium, 3 (7%) proliferative endometrium and 1 (3%) endometrial hyperplasia. Inter-rater agreement for the ultrasound diagnoses of uniformly thickened endometrium, polyp and

cancer, were 14/16 (87.5%), 22/30 (73.3%) and 28/34 (82.4%), respectively; the inter-rater reliability was good (κ = 0.69, 95% C.I. 0.49-0.88). When the ultrasound diagnoses were combined as either cancer or no cancer, the inter-rater agreement was 85% and the inter-rater reliability was also good (κ = 0.78, 95% C.I. 0.61-0.95). Rater A correctly identified 14/16 cancers and Rater B identified 15/16. Endometrial cancers were misdiagnosed as benign polyps on ultrasound in two women by Rater A, and in one woman by Rater B. The overall accuracies of Rater A and Rater B in differentiating between benign endometrial pathologies and malignancy were 90% and 85%, respectively.

Conclusions

Our results showed good inter-rater reliability of subjective pattern recognition in diagnosing uniformly thickened endometrium, polyp and cancer on ultrasound in women with postmenopausal bleeding. Our findings should facilitate wider use of subjective pattern recognition in routine clinical practice.

Introduction

Endometrial cancer (EC) is found in approximately 9% of women with postmenopausal bleeding (PMB)[1]. Measurement of the endometrial thickness (ET) can be used to triage women into low-risk or high risk-groups, with ET ≥3-5mm considered as high-risk of malignancy[2-6]. However, the specificity of ET in predicting EC is poor and therefore women with malignancy cannot be effectively prioritised for histological assessment.

Some studies have assessed whether the diagnostic accuracy of ultrasound for EC can be improved by assessing the endometrial morphological features and vascular pattern on Doppler ultrasound[7-15]. However, it was difficult to conclude from previous studies as different terminologies were used to describe the ultrasound findings. Therefore, the International Endometrial Tumor Analysis (IETA) group in 2010 published a consensus opinion on the terms, definitions and measurements of the endometrium[16]. Using the IETA terminologies, the most commonly reported features of EC are an endometrium with heterogeneous echogenicity, irregular or ill-defined endometrial-myometrial junction, and multiple vessels with focal or multifocal origins on Doppler ultrasound[17-19]. The presence of these ultrasound features may improve the diagnosis of EC, in addition to the measurement of ET. Subjective pattern recognition refers to the identification of these ultrasound features, in order to predict the presence or absence of EC. Recently, Dueholm et al.[20] reported that up to three-quarters of EC in women with PMB could potentially be identified by subjective pattern recognition and thereby fast-tracked for histological confirmation. Nevertheless, some are concerned about its reproducibility and suggested that it should be reserved for experts in centres of excellence only[21]. Indeed, the inter-rater reliability of ultrasound subjective pattern recognition and diagnostic accuracy may be poorer in less experienced operators[22-24]. Furthermore, the intra- and inter-rater reliabilities of using IETA terminology to describe ultrasound images of the

endometrium were also found to be poor[25]. Notwithstanding these potential limitations, others suggested that we should invest more in ultrasound training and consider limiting our use of terminologies to those that are most reproducible[26].

There are few reproducibility studies on ultrasound subjective pattern recognition in the literature. This study aimed to assess the inter-rater reliability of using subjective pattern recognition to diagnose EC in women with PMB, between an experienced and a less experienced operator.

Methods

This was a cross-sectional study conducted between Oct 2016 – Dec 2017 at a gynaecological rapid access clinic of a university teaching hospital. Patients were referred via their general practitioners urgently with a history of postmenopausal bleeding. Menopause was defined as amenorrhoea lasting ≥12 months in women over the age of 45. Consecutive patients who attended our clinic during the study period with an ET of ≥4.5mm on transvaginal ultrasound scan were included. We excluded women who were on hormone replacement therapy, tamoxifen or with a known history of gynaecological malignancy. If the endometrium could not be visualised on ultrasound assessment, women were offered either saline infusion sonography (SIS) or hysteroscopy, and they were excluded from the study because it was not practical to subject women to more than one SIS examination to assess the inter-rater reliability.

All women included in the study were examined by both Rater A and Rater B during a single outpatient visit. Rater A was a second-year clinical fellow in gynaecological ultrasound scan who had performed a total of approximately 2400 examinations before starting the study. This included assessments of over 200 women who presented with a history of postmenopausal bleeding. Rater B had over 30 years of experience in ultrasound and is a recognised expert in gynaecological ultrasound. The two raters performed their independent assessments in the absence of each other and they were blinded of each other's results. Each rater was allowed up to 10 mins to complete their assessment. The findings were recorded by an independent healthcare assistant who did not take part in the study.

All women were examined in the lithotomy position with an empty bladder after informed consent. The ultrasound assessments were performed systematically. First, the cervix and

uterine corpus were identified in the transverse plane. The uterine corpus was then assessed by examining a series of parallel scanning planes, starting from the internal cervical os to the top of the uterine fundus. The ET was measured using the "double-layer measurement technique"[16]. If there was fluid in the endometrial cavity, this was subtracted from the measurement of ET.

All women with an ET of ≥4.5mm were categorised into one of the following three groups based on endometrial morphological features on greyscale ultrasound and vascular patterns on colour Doppler that are in keeping with the IETA terminologies[16] (Figure 1):

1. Uniformly thickened endometrium – the endometrium appears uniform with no focal lesions, an intact midline echo and intact endometrial-myometrial junction. On Doppler ultrasound, it appears avascular or poorly vascularized (colour score ≤2).
2. Benign endometrial polyp – there is a well-defined localized lesion with a regular outline within the endometrial cavity. The surrounding endometrium appears regular with an intact endometrial-myometrial junction. On Doppler ultrasound, there is a single dominant vessel with or without branching, or there is no detectable vascularity.
3. Endometrial cancer – the endometrium appears heterogeneous or there is an irregular focal lesion. The endometrial-myometrial junction could be intact or it is interrupted, which is suggestive of myometrial invasion. On Doppler ultrasound, there are multiple vessels with focal or multifocal origins.

Endometrial sampling with pipelle biopsy was offered to all women with ultrasound diagnoses of EC or those with a uniformly thickened endometrium in our outpatient clinic. Women with endometrial polyps were offered hysteroscopic polypectomy and endometrial biopsy.

All ultrasound examinations were conducted with the same ultrasound equipment (Voluson E8, GE Healthcare Ultrasound, Milwaukee, WI, USA), equipped with a 4-9 MHz transvaginal probe.

Study outcomes

The primary outcome was the inter-rater reliability in the ultrasound diagnoses of uniformly thickened endometrium, polyp and cancer using subjective pattern recognition.

Statistical analysis

The first analysis was to determine the agreement and reliability between the two raters on the ultrasound diagnoses of uniformly thickened endometrium, polyp and cancer. Then a second analysis was performed with the categories combined as either EC or no endometrial cancer. Due to the categorical nature of the measurements, the reliability between raters was assessed using Cohen's kappa (κ) statistic, which represents the consistency of raters in their diagnoses. This is measured on a scale ranging up to a maximum agreement of 1. We adopted the classification in which a κ value of ≤0.20 is considered a very poor agreement between the two raters; 0.21-0.40 poor agreement; 0.41-0.60 moderate agreement; 0.61-0.80 good agreement; and 0.81-1.00 very good agreement[27].

Descriptive statistical methods were used to describe the study population. The statistical analyses were performed using IBM SPSS Statistics for Windows, version 25 (IBM Corp., Armonk, N.Y., USA).

Sample size calculation

This study aimed to determine the inter-rater reliability of subjective pattern recognition for the three possible ultrasound diagnoses of uniformly thickened endometrium, polyp and cancer in

women with postmenopausal bleeding. The minimum value for Cohen's kappa coefficient to be expected according to a previous study by Dueholm et al.[28] was 0.5 when there is assumed to be no agreement in the first place. When the power and alpha are pre-specified at 80% and 0.05%, respectively; a minimum sample of 18 women was required[29]. However, as the proportions of women in each of the three categories were not expected to be equal, we multiplied the minimum sample size by two to accommodate this variation. Hence, the required minimum sample size was 36 women (i.e. 18 x 2), which we have rounded up to include 40 women.

Ethical approval and reporting

Ethical approval was sought from the local research ethics committee, who approved the study but deemed that, as the ultrasound assessments in our study were not different to standard clinical practice, a full ethical approval was not required.

At our university teaching hospital, all ultrasound scans are performed by doctors in training who are under the supervision of specialist consultants. It is our routine practise that if a woman presents with postmenopausal bleeding and the endometrium is thickened, they are re-examined by an experienced consultant gynaecologist to verify their findings and advise on further investigations and management. We followed the Guidelines for Reporting Reliability and Agreement Studies (GRRAS statement) in the conduct and reporting of our research[27].

Results

There were 52 eligible women with PMB and ET of ≥4.5mm on transvaginal ultrasound scan during the study period. We excluded 12 women who were on hormone replacement therapy, tamoxifen or had a history of primary gynaecological malignancy. The remaining 40 women were included in the study. Their clinical characteristics and final histological diagnoses are presented in Table 1. A total of 16 patients underwent a hysterectomy, 17 had a hysteroscopy and 7 had outpatient pipelle endometrial biopsy.

Inter-rater agreement and reliability of ultrasound diagnoses by subjective pattern recognition

Ultrasound diagnoses of uniformly thickened endometrium, polyp and cancer by the two raters are shown in Table 2.

A total of 80 independent ultrasound diagnoses were made between the two raters. Agreements on uniformly thickened endometrium, polyp and cancer were 14/16 (87.5%), 22/30 (73.3%) and 28/34 (82.4%), respectively. Overall, the inter-rater reliability of the ultrasound diagnoses was good between the two raters with a kappa statistic of 0.69 (95% confidence interval (C.I.), 0.49-0.88).

When the three categories were dichotomised as either cancer or no cancer, in which the no cancer group included both women with uniformly thickened endometrium and polyps, the two raters were in agreement on 68/80 (85%) cases. Both raters agreed on the diagnosis of EC in 14 women but disagreed on 6 others. In 4 of these 6 women, Rater A diagnosed benign polyps, whereas Rater B diagnosed EC. In the other 2 cases, the opposite occurred. No cases of uniformly thickened endometrium were diagnosed when the other rater diagnosed EC. Overall, the inter-rater reliability on the presence or absence of EC was good with a kappa statistic of 0.78 (95% C.I. 0.61-0.95).

Diagnostic accuracy of ultrasound diagnoses by subjective pattern recognition

The final histological diagnoses against the ultrasound diagnoses of Rater A and Rater B are summarised in Tables 3 and 4, respectively. Rater A correctly identified 14/16 women with EC and Rater B identified 15/16. Rater A misdiagnosed two cases of EC as benign polyps, whereas Rater B misdiagnosed one case of EC as a benign polyp. Rater A had two false-positive diagnoses of EC, which were subsequently found to be a benign polyp and a disorganised proliferative endometrium in the final histological assessments. Rater B had three cases of false-positive diagnoses of EC, two of these were benign polyps and there was one with disorganised proliferative endometrium. The overall accuracies of Rater A and Rater B in differentiating between benign endometrial pathologies and malignancy were 90% and 85%, respectively.

Discussion

Our results showed good inter-rater reliability (κ = 0.69, 95% C.I. 0.49-0.88) in the ultrasound diagnoses of uniformly thickened endometrium, polyp and cancer by subjective pattern recognition, between an experienced operator and a less experienced operator. A kappa statistic of 0.69 meant that the raters accounted for 69% of the agreement over and above what would be expected by chance alone.

A previous study by Dueholm et al.[28] also investigated the inter-rater reliability of subjective pattern recognition in women with PMB and ET of ≥5mm. In their study, two experienced raters independently assessed 122 recorded videotapes of ultrasound examinations. Each woman was categorised into having no endometrial pathology, hyperplasia, benign polyp or cancer. Their reported inter-rater reliability was initially moderate (κ = 0.47, 95% C.I. 0.30-0.63). However, when the authors repeated their analysis by including videotapes of "high-quality" only, the inter-rater reliability was good (κ = 0.73, 95% C.I. 0.56-0.90) rather than moderate. We cannot easily compare our findings with those of Dueholm et al[28]; because, firstly, our study utilized real-time scanning rather than recorded videotapes. Secondly, we included only three possible ultrasound diagnoses rather than four, as we did not classify endometrial hyperplasia as a separate category; the reason for this is because we routinely offer women an endometrial biopsy in the clinic if they have a uniformly thickened endometrium. Therefore, an additional attempt to identify women who may have endometrial hyperplasia rather than a benign proliferative endometrium on ultrasound would not have made a clinical difference to their initial management. Furthermore, by reducing the number of available options to the rater, we aimed to improve the inter-rater reliability as concluded in another study by Sladkevicius et al[25].

Although reproducibility studies are important before a test can be introduced into clinical practice, a systematic review by Coelho et al.[30] found that only 14% of reproducibility studies were considered well designed and had interpreted their results appropriately. The authors criticised that most studies lacked independent acquisitions, blinded analyses and correct statistical analysis. In particular, 54.4% of reproducibility studies in ultrasound techniques in O&G were performed using only static images or video recordings acquired by a single rater. As a result, it is difficult to comment on how the inter-rater reliability could be affected by having different raters performing the ultrasound examinations. The strength of our study is that each rater performed their independent ultrasound examination and they were blinded to each other's findings. To the best of our knowledge, this is the first reported study where inter-rater reliability of subjective pattern recognition for EC has been tested with independent acquisitions.

Our study showed that the majority of women with EC can be identified on ultrasound by subjective pattern recognition and the inter-rater reliability on diagnosing EC was good (κ = 0.78, 95% C.I. 0.61-0.95). Both raters misdiagnosed EC as benign polyps only. As we routinely offer hysteroscopic polypectomy to all symptomatic women with suspected polyps, most EC is diagnosed subsequently following hysteroscopy. In view of the difficulty in distinguishing between benign polyps and polyps that harbour focal areas of premalignancy or malignancy, urgent hysteroscopic polypectomy is still warranted in all women who are symptomatic of postmenopausal bleeding. Further tests such as saline infusion sonography (SIS) may be helpful, as it may be easier to identify an irregular outline of malignant polyps[14]. More research is needed to improve the risk prediction of pre-malignant or malignant polyps in women with PMB.

Despite our small sample size, we showed that ultrasound diagnoses of uniformly thickened endometrium, polyp and cancer by subjective pattern recognition correlated well with the final histological diagnoses. This is encouraging because we could offer more individualised management according to the most likely diagnosis. In women with likely EC, they can be fast-tracked to a gynae-oncology centre following an outpatient endometrial biopsy. Whereas women with endometrial polyps on ultrasound can be offered an urgent hysteroscopy instead. Finally, those with a uniformly thickened endometrium without any suspicious ultrasound features for EC or polyp can be investigated less urgently.

We only included women with an ET of ≥4.5mm in our study because the prevalence of EC in women with an ET below this threshold is low. However, some EC may predominantly be confined to the myometrium, and therefore if we rely solely on the measurement of ET to exclude cancer, then some women with EC could potentially be misdiagnosed with a presumed "thin" endometrium[31]. Other studies have also reported that type II EC, such as uterine papillary serous carcinoma, clear cell carcinoma and other high-grade endometrial carcinomas are more likely to be found in women with a "thin" endometrium[32]. Future studies should consider the evaluation of subjective pattern recognition in women with PMB regardless of their ET.

There were several limitations to our study. Firstly, both raters belonged to the same academic unit, used the same high-end ultrasound machine and performed the ultrasound examinations in the same routine. These conditions may vary as ultrasound operators tend to use a wide range of machines and are trained to perform their examination in various routines. Secondly, it was not possible to blind the raters of the fact that they were taking part in a research study, therefore we cannot exclude the potential bias of the Hawthorn effect. Thirdly, we excluded all women with an unsatisfactory or suboptimal view of the endometrium as they were offered

SIS or hysteroscopy instead. Therefore, our results may not apply to women with a distorted endometrial cavity due to submucosal fibroids, severe adenomyosis or other uterine pathologies. Fourthly, our study included a 40% prevalence of EC which is higher than expected; given that a recent meta-analysis reported a pooled prevalence of only 19%[1]. This could be due to our small sample size, a higher prevalence of malignancy at a tertiary gynae-oncology centre or selection bias as we excluded many women with benign endometrial pathologies who underwent SIS.

In conclusion, there was good inter-rater reliability of diagnosing EC on ultrasound by subjective pattern recognition in women with PMB. The efficacy of subjective pattern recognition in fast-tracking women with EC for an earlier histological confirmation and referral to gynae-oncology centres should be assessed further in future studies.

References:

1. Clarke MA, Long BJ, Del Mar Morillo A, Arbyn M, Bakkum-Gamez JN, Wentzensen N. Association of Endometrial Cancer Risk With Postmenopausal Bleeding in Women. JAMA Internal Medicine. 2018;178(9):1210.
2. Karlsson B, Granberg S, Wikland M, Ylöstalo P, Torvid K, Marsal K, et al. Transvaginal ultrasonography of the endometrium in women with postmenopausal bleeding--a Nordic multicenter study. Am J Obstet Gynecol. 1995;172(5):1488-94.
3. Ferrazzi E, Torri V, Trio D, Zannoni E, Filiberto S, Dordoni D. Sonographic endometrial thickness: a useful test to predict atrophy in patients with postmenopausal bleeding. An Italian multicenter study. Ultrasound in Obstetrics and Gynecology. 1996;7(5):315-21.
4. Smith-Bindman R, Kerlikowske K, Feldstein VA, Subak L, Scheidler J, Segal M, et al. Endovaginal Ultrasound to Exclude Endometrial Cancer and Other Endometrial Abnormalities. JAMA. 1998;280(17).
5. Timmermans A, Opmeer BC, Khan KS, Bachmann LM, Epstein E, Clark TJ, et al. Endometrial thickness measurement for detecting endometrial cancer in women with postmenopausal bleeding: a systematic review and meta-analysis. Obstet Gynecol. 2010;116(1):160-7.
6. Wong A-W, Lao T-H, Cheung C, Yeung S, Fan H, Ng P, et al. Reappraisal of endometrial thickness for the detection of endometrial cancer in postmenopausal bleeding: a retrospective cohort study. BJOG: An International Journal of Obstetrics & Gynaecology. 2016;123(3):439-46.
7. Conoscenti G, Meir YJ, Fischer-Tamaro L, Maieron A, Natale R, D'Ottavio G, et al. Endometrial assessment by transvaginal sonography and histological findings after D & C in women with postmenopausal bleeding. Ultrasound in Obstetrics and Gynecology. 1995;6(2):108-15.
8. Weigel M, Friese K, Strittmatter HJ, Melchert F. Measuring the thickness-is that all we have to do for sonographic assessment of endometrium in postmenopausal women? Ultrasound in Obstetrics and Gynecology. 1995;6(2):97-102.
9. Weber G, Mere E, Bahlmanw E, Riisch B. Evaluation of different transvaginal sonographic diagnostic parameters in women with postmenopausal bleeding. Ultrasound in Obstetrics and Gynecology. 1998;12(4):265-70.
10. Sheikh M, Sawhney S, Khurana A, Al-Yatama M. Alteration of sonographic texture of the endometrium in post-menopausal bleeding. Acta Obstetricia et Gynecologica Scandinavica. 2000;79(11):1006-10.
11. Randelzhofer B, Prömpeler HJ, Sauerbrei W, Madjar H, Emons G. Value of sonomorphological criteria of the endometrium in women with postmenopausal bleeding: a multivariate analysis. Ultrasound in Obstetrics and Gynecology. 2002;19(1):62-8.
12. Epstein E, Ramirez A, Skoog L, Valentin L. Transvaginal sonography, saline contrast sonohysterography and hysteroscopy for the investigation of women with postmenopausal bleeding and endometrium > 5 mm. Ultrasound in Obstetrics and Gynecology. 2001;18(2):157-62.
13. Alcázar JL, Castillo G, Mínguez JÁ, Galán MJ. Endometrial blood flow mapping using transvaginal power Doppler sonography in women with postmenopausal bleeding and thickened endometrium. Ultrasound in Obstetrics & Gynecology. 2003;21(6):583-8.
14. Epstein E, Valentin L. Gray-scale ultrasound morphology in the presence or absence of intrauterine fluid and vascularity as assessed by color Doppler for discrimination between

benign and malignant endometrium in women with postmenopausal bleeding. 2006;28(1):89-95.
15. Opolskiene G, Sladkevicius P, Valentin L. Ultrasound assessment of endometrial morphology and vascularity to predict endometrial malignancy in women with postmenopausal bleeding and sonographic endometrial thickness ≥ 4.5 mm. Ultrasound in Obstetrics and Gynecology. 2007;30(3):332-40.
16. Leone FPG, Timmerman D, Bourne T, Valentin L, Epstein E, Goldstein SR, et al. Terms, definitions and measurements to describe the sonographic features of the endometrium and intrauterine lesions: a consensus opinion from the International Endometrial Tumor Analysis (IETA) group. Ultrasound in Obstetrics and Gynecology. 2010;35(1):103-12.
17. Kabil Kucur S, Temizkan O, Atis A, Gozukara I, Uludag EU, Agar S, et al. Role of endometrial power Doppler ultrasound using the international endometrial tumor analysis group classification in predicting intrauterine pathology. Arch Gynecol Obstet. 2013;288(3):649-54.
18. Madkour NM. An ultrasound risk-scoring model for prediction of endometrial cancer in post-menopausal women (using IETA terminology). Middle East Fertility Society Journal. 2017;22(3):201-5.
19. Epstein E, Fischerova D, Valentin L, Testa AC, Franchi D, Sladkevicius P, et al. Ultrasound characteristics of endometrial cancer as defined by International Endometrial Tumor Analysis (IETA) consensus nomenclature: prospective multicenter study. Ultrasound in Obstetrics & Gynecology. 2018;51(6):818-28.
20. Dueholm M, Marinovskij E, Hansen ES, Møller C, Ørtoft G. Diagnostic methods for fast-track identification of endometrial cancer in women with postmenopausal bleeding and endometrial thickness greater than 5 mm. Menopause. 2015;22(6):616-26.
21. Angioli R, Aloisi A, Capriglione S, Plotti F. Numquam ponenda est pluralitas sine necessitate? Ultrasound in Obstetrics & Gynecology. 2014;44(3):372-3.
22. Eriksson LSE, Lindqvist PG, Flöter Rådestad A, Dueholm M, Fischerova D, Franchi D, et al. Transvaginal ultrasound assessment of myometrial and cervical stromal invasion in women with endometrial cancer: interobserver reproducibility among ultrasound experts and gynecologists. Ultrasound in Obstetrics & Gynecology. 2015;45(4):476-82.
23. Green RW, Valentin L, Alcazar JL, Chiappa V, Erdodi B, Franchi D, et al. Endometrial cancer off-line staging using two-dimensional transvaginal ultrasound and three-dimensional volume contrast imaging: Intermethod agreement, interrater reliability and diagnostic accuracy. Gynecol Oncol. 2018;150(3):438-45.
24. Alcázar JL, Ajossa S, Floris S, Bargellini R, Gerada M, Guerriero S. Reproducibility of Endometrial Vascular Patterns in Endometrial Disease as Assessed by Transvaginal Power Doppler Sonography in Women With Postmenopausal Bleeding. Journal of Ultrasound in Medicine. 2006;25(2):159-63.
25. Sladkevicius P, Installé A, Van Den Bosch T, Timmerman D, Benacerraf B, Jokubkiene L, et al. International Endometrial Tumor Analysis (IETA) terminology in women with postmenopausal bleeding and sonographic endometrial thickness ≥ 4.5 mm: agreement and reliability study. Ultrasound in Obstetrics & Gynecology. 2018;51(2):259-68.
26. Neto MC. Re: International Endometrial Tumor Analysis (IETA) terminology in women with postmenopausal bleeding and sonographic endometrial thickness ≥ 4.5 mm: agreement and reliability study. P. Sladkevicius, A. Installé, T. Van den Bosch, D. Timmerman, B. Benacer. Ultrasound in Obstetrics & Gynecology. 2018;51(2):167-8.

27. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. Journal of Clinical Epidemiology. 2011;64(1):96-106.
28. Dueholm M, Hjorth IM, Secher P, Jørgensen A, Ørtoft G. Reproducibility of Endometrial Pathologic Findings Obtained on Hysteroscopy, Transvaginal Sonography, and Gel Infusion Sonography in Women With Postmenopausal Bleeding. J Minim Invasive Gynecol. 2015;22(6):1036-44.
29. Bujag MA BN. Guidelines of the minimum sample size requirements for Cohen's Kappa. Epidemiology Biostatistics and Public Health. 2017;14(2).
30. Coelho Neto MA, Roncato P, Nastri CO, Martins WP. True Reproducibility of UltraSound Techniques (TRUST): systematic review of reliability studies in obstetrics and gynecology. Ultrasound in Obstetrics & Gynecology. 2015;46(1):14-20.
31. Naftalin J, Nunes N, Hoo W, Arora R, Jurkovic D. Endometrial cancer and ultrasound: why measuring endometrial thickness is sometimes not enough. Ultrasound in Obstetrics & Gynecology. 2012;39(1):106-9.
32. Wang J, Wieslander C, Hansen G, Cass I, Vasilev S, Holschneider CH. Thin endometrial echo complex on ultrasound does not reliably exclude type 2 endometrial cancers. Gynecologic Oncology. 2006;101(1):120-5.

Figure legends

Figure 1 - A. uniformly thickened endometrium

Figure 1 - B. benign endometrial polyp

Figure 1 - C. endometrial cancer

Table 1. Patient characteristics and final histological diagnoses (n=40)	
Characteristics	Value
Age (years)[2]	61 (57-69)
Time since menopause (years)[2]	9.5 (5.0-19.5)
Nulliparity[2]	12 (30)
Caucasian ethnicity[2]	30 (75)
BMI (kg/m^2)[1]	29.3 (24.2-34.4)
Endometrial thickness (mm)[1]	11.0 (6.2-20.3)
Uterine fibroids[2]	19 (48)
Adenomyosis[2]	11 (28)
Endometrial cancer[2]	16 (40)
- Stage IA	6 (37.5)
- Stage IB	4 (25)
- Stage II	0 (0)
- Stage IIIA	1 (6.3)
- Stage IIIB	1 (6.3)
- Stage IIIC1	2 (12.5)
- Stage IIIC2	1 (6.3)
- Stage IV	1 (6.3)
Histological subtype	
Endometrioid[2]	10 (62.5)
- Grade 1	5 (50)
- Grade 2	3 (30)
- Grade 3	2 (20)
Non-endometrioid[2]	6 (37.5)
- Serous	2 (33.3)
- Carcinosarcoma	2 (33.3)
- Undifferentiated	1 (16.7)
- Neuroendocrine	1 (16.7)
Benign endometrial pathologies[2]	24 (60)
- Endometrial polyp	16 (66.7)
- Atrophic endometrium	4 (16.7)
- Proliferative endometrium	3 (12.5)
- Endometrial hyperplasia	1 (4.2)
Results are presented as median (interquartile range)[1] or n (%)[2]	

Table 2. Ultrasound diagnoses by both raters (n=40)					
	Rater B				
Rater A		Uniformly thickened endometrium	Polyp	Cancer	**Total**
	Uniformly thickened endometrium	7	2	0	9
	Polyp	0	11	4	15
	Cancer	0	2	14	16
	Total	7	15	18	40

Table 3. Ultrasound diagnoses by Rater A and the final histological diagnoses (n=40)

		Ultrasound diagnoses by Rater A			
		Uniformly thickened endometrium	Polyp	Cancer	**Total**
Histological diagnoses	Atrophic endometrium	4	0	0	4
	Disorganised proliferative endometrium	2	0	1	3
	Endometrial hyperplasia	1	0	0	1
	Benign polyp	2	13	1	16
	Cancer	0	2	14	16
	Total	9	15	16	40

Table 4. Ultrasound diagnoses by Rater B and the final histological diagnoses (n=40)					
	Ultrasound diagnoses by Rater B				
		Uniformly thickened endometrium	Polyp	Cancer	**Total**
Histological diagnoses	Atrophic endometrium	3	1	0	4
	Disorganised proliferative endometrium	2	0	1	3
	Endometrial hyperplasia	1	0	0	1
	Benign polyp	1	13	2	16
	Cancer	0	1	15	16
	Total	7	15	18	40

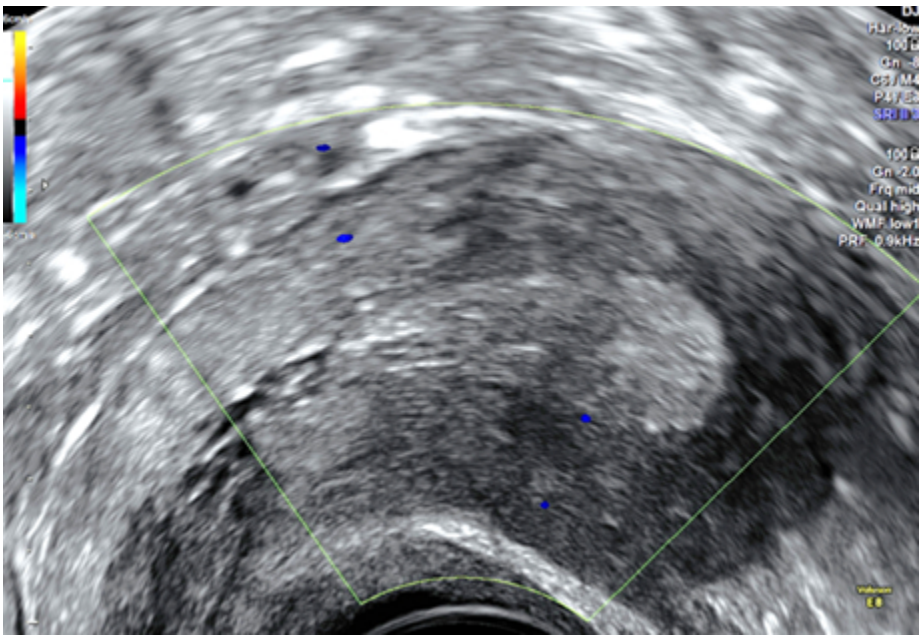

Figure 1 - A. uniformly thickened endometrium

121x83mm (96 x 96 DPI)

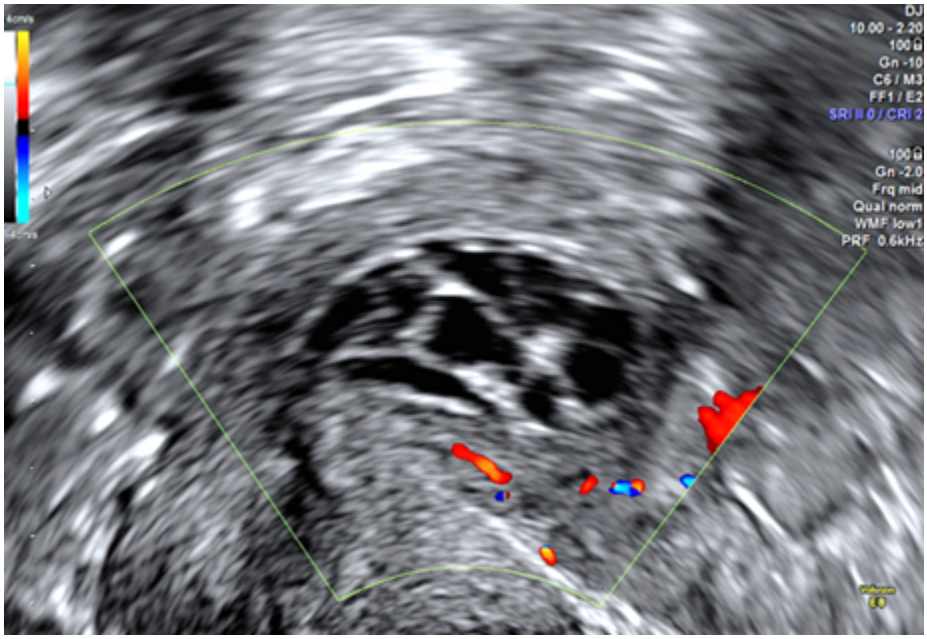

Figure 1 - B. benign endometrial polyp

121x83mm (96 x 96 DPI)

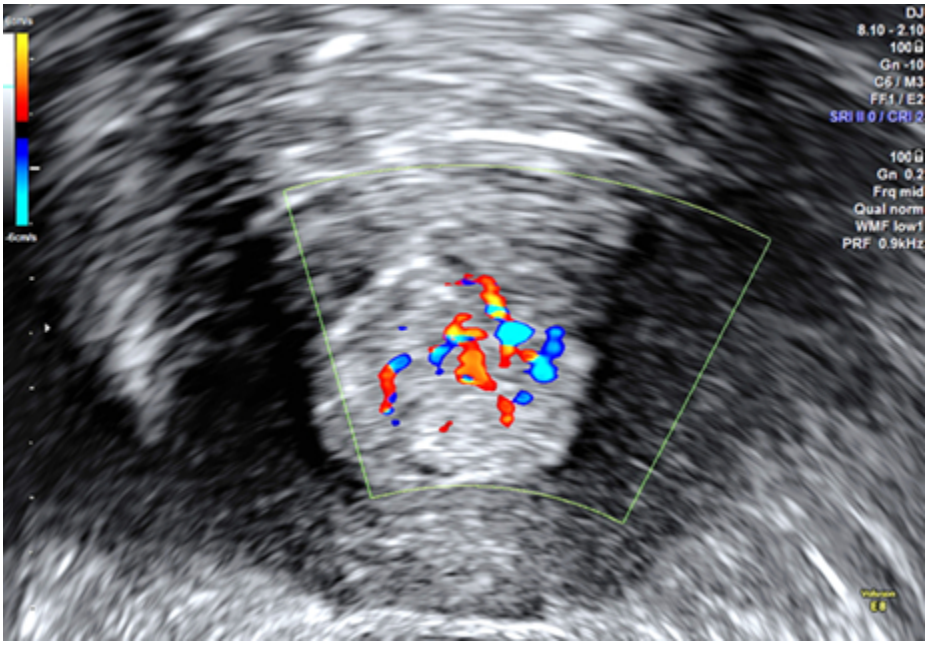

Figure 1 - C. endometrial cancer

121x84mm (96 x 96 DPI)