Archives of Gynecology and Obstetrics (2019) 299:779–789
https://doi.org/10.1007/s00404-019-05050-x

GENERAL GYNECOLOGY

Three-dimensional ultrasound in the diagnosis and the classification of congenital uterine anomalies using the ESHRE/ESGE classification: a diagnostic accuracy study

Anna Kougioumtsidou[1] · Themistoklis Mikos[1] · Grigoris F. Grimbizis[1] · Aikaterini Karavida[1] · Theodoros D. Theodoridis[1] · Alexandros Sotiriadis[1] · Basil C. Tarlatzis[1] · Apostolos P. Athanasiadis[1]

Received: 7 October 2018 / Accepted: 5 January 2019 / Published online: 19 January 2019

Abstract

Study objective To estimate the diagnostic accuracy of three-dimensional ultrasonography (3D US) compared to hysteroscopy/laparoscopy, in the investigation of uterine congenital anomalies using the ESHRE/ESGE classification of female genital tract congenital anomalies.

Design Prospective blind, comparative, cohort study.

Setting University Tertiary Hospital and affiliated private Hospital.

Patients and methods Sixty-two women consecutively referred with a suspected diagnosis of uterine congenital anomalies. The ESHRE/ESGE classification of congenital anomalies of the female genital tract was used for the description of abnormal findings.

Interventions All patients underwent (1) 3D US and (2) hysteroscopy with laparoscopy to establish the final diagnosis.

Results Concordance between 3D US and hysteroscopy with laparoscopy about the type and the classification of uterine anomaly was verified in 61 cases, including all those with septate uterus, dysmorphic uterus, bicorporeal, hemi-uterus or unicorporeal, and aplastic uterus and one out of two with normal uterus. For the diagnosis of septate uteri, which was the most common anomaly, the sensitivity of 3D US was 100%, the specificity was 92.3%, the PPV was 98% and the NPV was 100%, with kappa index 0.950. For bicorporeal, dysmorphic uterus, hemi-uteri or unicorporeal and aplastic uterus the sensitivity, specificity, PPV and NPV were all 100% with $K = 1.00$. Overall, 3D US showed perfect diagnostic accuracy (Kappa index = 0.945) in the detection of congenital uterine anomalies.

Conclusion 3D US appears to be a very accurate method for the diagnosis of congenital uterine anomalies.

Keywords Three-dimensional ultrasound · Congenital anomalies of the female genital tract · Hysteroscopy · Laparoscopy · ESHRE/ESGE classification system

Introduction

Female genital anomalies are the result of the abnormal formation, canalization and/or fusion of the paramesonephric ducts and/or from the defective absorption of the midline septum [1]. The total prevalence of female genital tract malformations is estimated to be 5.5–6.7% for the general population, 7.3–8.0% for the infertile population, and 13.3–16.7% for the recurrent pregnancy loss population [2, 3]. Septate uterus is the most common congenital uterine malformation in high-risk populations [3].

Accurate diagnosis and classification of the congenital uterine anomalies are mandatory, especially in the cases with septate and bicorporeal uterus, where the patients may

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00404-019-05050-x) contains supplementary material, which is available to authorized users.

✉ Anna Kougioumtsidou
akougioum@gmail.com

[1] 1st Department of Obstetrics and Gynecology, Papageorgiou General Hospital, Ring Road, Municipality Pavlos Melas, N. Efkarpia, 56403 Thessaloniki, Greece

benefit from interventional therapeutic approaches. There have been described several classifications of uterine malformations [1]. Until 2013, three systems have been proposed for the classification of the female genital tract anomalies: (1) the American Fertility Society's (AFS) based on the previous work of Buttram and Gibbons [4, 5], (2) the embryological-clinical classification system of genito-urinary malformations [6], and (3) the Vagina, Cervix, Uterus, Adnexae and associated Malformations system, based on the tumor nodes metastases (TNM) principle in oncology [7].

More recently, the European Society of Human Reproduction and Embryology (ESHRE) and the European Society for Gynecological Endoscopy (ESGE) have published the new ESHRE/ESGE classification system of congenital anomalies of the female genital tract, where six main classes of uterine anomalies were proposed; normal uterus was adopted as class U0, dysmorphic uterus as class U1, septate uterus as class U2, bicorporeal uterus as class U3, unicorporeal or hemi-uterus as class U4, and aplastic uterus as class U5; there is still the possibility of potentially unclassified cases to be grouped in as class U6. This system uses uterine anatomy as the basic characteristic for the design of the main classes; main sub-classes are also based on different degrees of uterine deformity that may have clinical significance. Cervical and vaginal anomalies are classified in independent co-existent sub-classes, giving clarity in the anatomical representation of each anomaly [8, 9].

Currently, a lot of innovative, non-invasive, high accuracy diagnostic techniques are available allowing for the accurate imaging of uterine anatomy. Two-dimensional ultrasonography (2D US) is the most available diagnostic tool and it is routinely used as the first approach in diagnosis of the uterine anomalies in anomalies in "asymptomatic" women [10]. 2D US has a reported mean accuracy in diagnosing congenital uterine anomalies of 86.6% [10–12]. However, the distinction between the different types of anomalies is difficult with 2D US [2, 13]. Magnetic Resonance Imaging (MRI) has been proposed as a reliable diagnostic approach with a mean accuracy of 85.8%, especially in the cases with suspected complex and obstructing anomalies, but it is more expensive and less available than ultrasound [10]. Three-dimensional ultrasonography (3D US) is a diagnostic method with reported high accuracy in the diagnosis of uterine anomalies reaching [14–18]; the mean diagnostic accuracy was 97.6% (CI 94.3–100) [10]. 3D US is a non-invasive, reproducible, and less costly, compared to the MRI, procedure that provides information about both the internal and external contour of the uterus by displaying the coronal plane of the uterus. Endoscopic evaluation was considered until recently as "the gold standard" in the confirmation of diagnosis of female genital anomalies; however, 3D US due to its high diagnostic accuracy has been proposed as non-invasive "gold standard" approach for these malformations [2, 3, 10].

The purpose of this study was to evaluate the accuracy of 3D US in the diagnosis of Mullerian anomalies compared to Hysteroscopy and Laparoscopy, using the ESHRE/ESGE classification of congenital anomalies of the female genital tract.

Materials and methods

This is a prospective, blind, comparative study, performed in a tertiary referral setting from 2012 to 2016. Ethical approval was as appropriately obtained (AUTH Protocol No. 49 28/06/2013). The trial was registered with the ISRCTN (No. 12592528).

The study design included women with suspected uterine malformations. The inclusion criterion for the study was a presumptive 2D US diagnosis of any uterine anomaly. Exclusion criteria were (1) pre-pubertal adolescents, (2) pregnancy, (3) menopause, and (4) any presence of uterine fibroids. The presenting symptoms were: (1) primary amenorrhea, (2) history of infertility, (3) recurrent abortions, and (4) suspicious findings for presence of a congenital anomaly in 2D US.

All women who consecutively presented in our clinic with suspicion of congenital uterine anomalies in the 2D US were invited to participate in the study.

All participants had (1) a medical interview and clinical examination, (2) 2D US, (3) 3D US to diagnose and classify any uterine malformations according to ESHRE/ESGE classification of congenital anomalies of female genital tract, and (4) hysteroscopy and laparoscopy to diagnose and classify any uterine malformations according to ESGE/ESHRE classification.

Three-dimensional ultrasonography (3D US)

All patients underwent transvaginal 3D US for assessment of the uterine anomalies during the luteal phase of their cycles, except for two cases with intact hymen and vaginal aplasia who had a transabdominal 3D scan. All examinations were performed by two gynecologists (A.A., A.K.) with extensive experience in gynecological ultrasound; the ultrasound systems Voluson 730 Expert (GE Medical Systems, Zipf, Austria) and Voluson E8 Expert (GE Medical Systems, Zipf, Austria) with volumetric transvaginal probe (5–9 MHz frequency) and volumetric convex transabdominal probes (6–12 MHz frequency) were used.

A standardized protocol for the description of the ultrasound findings is presented in Table 1.

Table 1 A standardized protocol for the description of the ultrasound findings

The patients were scanned with an empty bladder. According to the protocol, the uterus was initially scanned by 2D US from the cervix to the fundus in longitudinal and transverse sections. After having performed a standard 2D evaluation, a 3D volume of the uterus was acquired. The three-dimensional acquisition box was adjusted to include a clear area around the uterus to allow delineation of the myometrium. The uterus was visualized in the sagittal plane so that the endocervical canal and endometrial outline could be seen in their entirety. Automatic volume acquisition was performed using a sweep angle between 90°- and 120°-degrees to obtain a multiplanar view of the uterus. This rendering modality allows analysis of the volumes in three orthogonal planes and obtains the coronal view showing the external and fundal contours of the uterus. In cases where the uterus could not be included in a single volume, two volumes were acquired to allow the study of the whole uterine cavity. Additionally, when the transverse uterine distance was increased, the volume was obtained from the coronal plane. The uterine structure was studied on the basis of coronal view information and with uterine wall thickness used as reference point (uterine wall thickness was defined as the distance between the interostial line and external uterine profile at the mid-coronal plane of the uterus). 3D US volumes were then analyzed by the same clinicians (A.A., A.K.) to classify any uterine anomalies according to the ESHRE/ESGE classification of congenital anomalies of female genital tract

Hysteroscopy and laparoscopy

The final diagnosis for uterine anomaly was ascertained by hysteroscopy and laparoscopy (H & L). All the procedures were conducted by senior gynecologist (G.G.) with special interest in the diagnosis and treatment of congenital uterine anomalies. The clinician who performed the endoscopic procedures was blind to 3D US results.

A standardized protocol for the description of the endoscopic findings was used including estimation of pelvic and uterine anatomy as defined in the ESHRE/ESGE classification classes.

The vaginal and cervical anatomy was carefully inspected with gynaecological examination. Then, saline, 5-mm, diagnostic hysteroscopy was performed using the atraumatic vaginoscopic approach. In cases of double cervices, hysteroscopic evaluation of both cavities was done. By turning the endoscope around its longitudinal axis, a complete visualization of the cavity was achieved. According to the study protocol, the following uterine segments were assessed: (1) endocervical canal (shape, size, contour, length, presence of lesions or septum), (2) endometrial cavity (shape, size, contour, number, presence and description of endometrial lesions, presence and length of uterine septum), and (3) tubal ostia (place, patency). In cases of the presence of a fundal indentation, the surgeon estimated in cm the length of the septum. During laparoscopy, the size and the outer contour of the uterus were evaluated. In the cases with the intercornual cleft, its measurement was estimated in cm by the surgeon.

In cases of abnormal findings, operative endoscopy was performed at the same session. The cases of septate and T-shaped uterus underwent operative hysteroscopic assessment and treatment. All cases of bicorporeal, aplastic and hemi-uteri were confirmed by laparoscopy.

Description of the anatomic anomalies at the 3D US

The ESHRE/ESGE classification of female congenital abnormalities was used for the description of the anatomic uterine anomalies [8, 9] and is presented in Table 2.

Statistical analysis

Data were expressed as mean ± standard deviation (SD) for quantitative variables and as percentages for qualitative variables.

The diagnostic results of the 3D US for each class of uterine anomaly were compared to the diagnostic H & L results to determine the sensitivity, specificity, PPV, NPV and accuracy of 3D US.

The agreement between the 3-D ultrasonography and hysteroscopy/laparoscopy was estimated with linear weighted Kappa index.

The interobserver agreement for the diagnosis of each uterine anomaly was assessed by the linear weighted Kappa index. The interobserver agreement for the quantitative measurements was assessed by estimating the inter-class correlation coefficient (ICC), the comparison of two measurements using the paired samples t test and the Bland–Altman plot using a visual method of assessing observers agreement.

All tests were two sided; a p value of < 0.05 was used to denote statistical significance. All analyses were carried out using the statistical package SPSS vr 17.00 (Statistical Package for the Social Science, SPSS Inc., Chicago, IL, USA).

Results

From the 64 women who were initially included in the study, 62 patients participated in the final analysis. Two patients were excluded because of detection of uterine fibroids at 2D US. All 62 women with suspected 2D US diagnosis of uterine anomaly underwent 3D US and hysteroscopy/laparoscopy.

Table 2 Description of the anatomic uterine anomalies using the ESHRE/ESGE classification of female genital tract anomalies

According to this classification and with the aid of 3D US, the anomalies were classified as follows

Class U0 (normal uterus)

Class U1 or dysmorphic uterus (all cases with normal uterine outline but with an abnormal shape of the uterine cavity excluding septa); Class U1a—T-shaped uterus (narrow uterine cavity due to thickened lateral walls with a correlation 2/3 uterine corpus and 1/3 cervix); Class U1b (uterus infantilis); Class U1c—(all minor deformities of the uterine cavity)

Class U2 or septate uterus; Class U2a (partial septum) (Suppl. Fig. 1); Class U2b (Complete septum) (Suppl. Fig. 2). We used the distance between the interostial line and a parallel line on the top of the uterine fundus (i.e., the uterine wall thickness) and the distance between the interostial line and a parallel line on the top of midline indentation (i.e., the depth of the septum). To classify the cases with uterine septum, we used the proportion between these two measurements [9]. An internal indentation > 50% of the uterine wall thickness was used as a criterion for septate uterus; in these cases, the external contour should be straight or with indentation less then 50%

Class U3 or bicorporeal uterus; Class U3a (partial indentation); Class U3b (complete indentation); Class U3c (bicorporeal septate uterus). We used the distance between the interostial line and a parallel line joining the external outline of the uterine horns (i.e., the uterine wall) and the distance between the interostial line and a parallel line on the top of the uterine indentation (i.e., the depth of the indentation). To classify the cases with bicorporeal uteri, we used the proportion between these two measurements. The external indentation more than 50% of the uterine wall thickness was used as a criterion for Class U3a; the presence of external indentation that reached the internal cervical os was used as criterion for Class U3b; the presence of uterine septum > 150% of the uterine wall thickness along with the presence of external uterine indentation was used as criterion for Class U3c

Class U4 or hemi-uterus (Suppl. Fig. 3); Class U4a with rudimentary cavity; Class U4b without rudimentary cavity. The sonographic presence of a single uterine cavity and the demonstration of a single interstitial portion of Fallopian tube at the coronal plane was used as criteria for Class U4

Class U5 incorporates cases of uterine aplasia

Class U6 is kept for still unclassified cases

Co-existent cervical anomalies

***Sub-class C0 or normal cervix* incorporates all cases of normal cervical development**

Sub-class C1 (*Suppl. Fig.* 4) *or septate cervix is characterized by the presence of a normal externally rounded cervix with the presence of a septum*

Sub-class C2 or double cervix is characterized by the presence of two distinct externally rounded cervices; these two cervices could be either fully divided or partially fused. It could be combined with a complete bicorporeal uterus as a Class U3b/C2 in the formerly didelphys uterus

Sub-class C3 or unilateral cervical aplasia incorporates all cases of unilateral cervical formation. *It is characterized by the unilateral, only, cervical development; the contralateral part could be either incompletely formed or absent.* Obviously, this has happened in Class U4 patients; however, this is not necessary to be mentioned in the final classification report (Class U4 instead of Class U4/C3) as being apparent. On the other hand, this sub-class gives the opportunity to classify other seldom anomalies such as complete bicorporeal uterus with unilateral cervical aplasia as Class U3b/C3, which is a severe obstructing anomaly

Sub-class C4 or cervical aplasia incorporates all cases of complete cervical aplasia but, also, those of severe cervical formation defects. *It is characterized either by the absolute absence of any cervical tissue or by the presence of severely defected cervical tissue such as cervical cord, cervical obstruction and cervical fragmentation.* The decision to include all variants of cervical dysgenesis in sub-class C4 was made in order to avoid an extremely extensive sub-classification, which does not seem to be user friendly. This sub-class could be combined with a normal or a defected uterine body and gives the opportunity to classify all obstructing anomalies due to cervical defects

Co-existent vaginal anomalies

Sub-class V0 or normal vagina incorporates all cases of normal vaginal development

Sub-class V1 or longitudinal non-obstructing vaginal septum. The incorporated anomaly in this sub-class is clear; it gives the opportunity to classify variants of septate or bicorporeal uteri together with septate or double cervices

Sub-class V2 or longitudinal obstructing vaginal septum. The incorporated anomaly in this sub-class is also clear and, its utility for the effective classification of obstructing anomalies due to vaginal defects is obvious

Sub-class V3 or transverse vaginal septum and/or imperforate hymen. This sub-class incorporates obviously different vaginal anomalies and their variants (mainly those of transverse vaginal septa); this was decided in order to avoid an extremely extensive sub-classification for the classification system's simplicity. The decision to put together those vaginal anomalies in this sub-class is due to the fact that they are usually present as isolated vaginal defects and they have the same clinical presentation (obstructing anomalies)

Sub-class V4 or vaginal aplasia incorporates all cases *of complete or partial vaginal aplasia*

Table 3 Demographic and clinical characteristics of the patients

Age (years); mean ± SD (min–max)	31.7 ± 5.6 (13–42)
Infertility/recurrent abortions; *n* (%)	33/62 (53.2%)
Number of pregnancies (0-1-2-3+); *n* (%)	26 (42%)/13 (21%)/18 (29%)/5 (8%)
Number of deliveries (0-1-2); *n* (%)	44 (71%)/11 (17%)/7 (11.3%)

Patients' clinical details

The demographic and clinical characteristics of the patients are shown in Table 3. There were 2/62 (3.2%) patients with primary amenorrhea, 33/62 (53.2%) with history of infertility and recurrent abortion, and 27/62 (43.6%) without any clinical manifestations. The mean age of women was 31.8 (SD 5.6).

H & L findings

The findings of H & L were: Normal uterus (Class U0) in 2/62 (3.2%) patient, T-shaped uterus (Class U1a) in 3/62 (4.8%) patients, septate uterus (Class U2) in 49/62 (78.9%) patients [30 (61.2%) patients with partial septum Class U2a and 19 (38.8%) patients with complete septum Class U2b], bicorporeal uterus (Class U3) in 3/62 (4.8%) patients, hemi-uterus (Class U4) in 4/62 (6.5%) patients, and aplastic uterus (Class U5) in 1/62 (1.6%) patient. The most common anomaly was septate uterus.

Twelve patients had also co-existent vaginal and cervical anomalies: 1 (1.6%) women with vaginal aplasia (sub-Class V4), 4 (6.4%) patients with longitudinal non-obstructing vaginal septum (sub-Class V1) and 2 (3.2%) women with transverse vaginal septum (sub-Class V3). Three (4.8%) women were diagnosed with double "normal" cervix (sub-Class C2), 1 (1.6%) with cervical aplasia (sub-Class C4) and 3 (4.8%) with septate cervix (sub-Class C1).

3D US interobserver variability

The interobserver reproducibility of the diagnosis of congenital uterine anomalies using 3D US was examined in all cases of the study. Moreover, in all cases where septate uterus was diagnosed with the 3D US, the reproducibility of the measurements of (1) the uterine wall thickness, (2) the length of internal indentation, and (3) the interostial distance were also examined. There was perfect agreement between the two observers in the diagnosis and the classification of uterine anomalies [kappa 0.945; 95% CI (0.84–1.00)] (Table 4). However, a case of normal uterus was diagnosed by Observer B as a septate uterus.

There was perfect agreement between the two operators in relation to quantitative variables based on intra-class correlation coefficients (ICC > 0.934) (Table 5). There is no statistical significant difference between observers for Uterine Wall Thickness ($p = 0.266$), Length of Internal Indentation ($p = 0.585$) and Interostial Distance ($p = 0.170$) (Table 6). Bland–Altman plots of interobserver reproducibility for

Table 4 Interobserver reliability of 3D US diagnosis of congenital uterine anomalies

Observer A	Observer B						
	U0	U1	U2	U3	U4	U5	Total
U0							
N	1	0	0	0	0	0	1
%	1.6	0.0	0.0	0.0	0.0	0.0	1.6
U1							
N	0	3	0	0	0	0	3
%	0.0	4.8	0.0	0.0	0.0	0.0	4.8
U2							
N	1	0	49	0	0	0	50
%	1.6	0.0	80.6	0.0	0.0	0.0	80.6
U3							
N	0	0	0	3	0	0	3
%	0.0	0.0	0.0	4.8	0.0	0.0	4.8
U4							
N	0	0	0	0	4	0	4
%	0.0	0.0	0.0	0.0	6.4	0.0	6.4
U5							
N	0	0	0	0	0	1	1
%	0.0	0.0	0.0	0.0	0.0	1.6	1.6
Total							
N	2	3	49	3	4	1	62
%	3.2	4.8	78.9	4.8	6.4	1.6	100.0

U0, normal uterus; U1, dysmorphic uterus; U2, septate uterus; U3, bicorporeal uterus; U4, hemi-uterus; U5 aplastic uterus

Table 5 Intra-class correlation coefficients between observers

	ICC (95% CI)	*p* Value
Uterine wall thickness	0.934 (0.89–0.96)	< 0.001
Length of internal indentation	0.994 (0.99–1.00)	< 0.001
Interostial distance	0.970 (0.95–0.98)	< 0.001

There is perfect agreement between observers in relation to quantitative variables based on intra-class correlation coefficients, uterine wall thickness [0.934; 95% CI (0.89–0.96)], length of internal indentation [0.994; 95% CI (0.99–1.00)], interostial distance [0.970; 95% CI (0.95–0.98)]

all measurements are shown in Figs. 1, 2 and 3. Inspection of scattergrams showed that > 95% of differences were within mean difference ± 2 SDs, thus confirming the agreement of observers. Readings by Observer A were on average 0.108–0.27 mm lower than measurements taken by Observer B. However, this small bias is unlikely to be of clinical significance.

3D US findings

The findings of 3D US were: Normal uterus (Class U0) in 1/62 (1.6%) patient, T-shaped uterus (Class U1a) in 3/62 (4.8%) patients, septate uterus (Class U2) in 50/62 (80.5%) patients (31 (62%) patients with partial septum (Class U2a) and 19 (38%) patients with complete septum (Class U2b), bicorporeal uterus (Class U3) in 3/62 (4.8%) patients,

Table 6 Comparison of quantitative variables between observers

	Uterine wall thickness	Length of internal indentation	Interostial distance
Observer A (*n* = 50)	9.24 ± 2.12	15.17 ± 9.24	33.34 ± 5.59
Observer B (*n* = 50)	9.36 ± 2.03	15.25 ± 9.33	33.61 ± 5.49
Mean difference ± SD	− 0.12 ± 0.75	− 0.08 ± 1.05	− 0.27 ± 1.36
95% CI	(− 0.33, 0.09)	(− 0.38, 0.22)	(− 0.65, 0.12)
t value (1.49)	− 1.124	0.550	− 1.394
p value	0.266	0.585	0.170

There is no statistical significant difference between observers for uterine wall thickness ($p = 0.266$), length of internal indentation ($p = 0.585$) and interostial distance ($p = 0.170$)

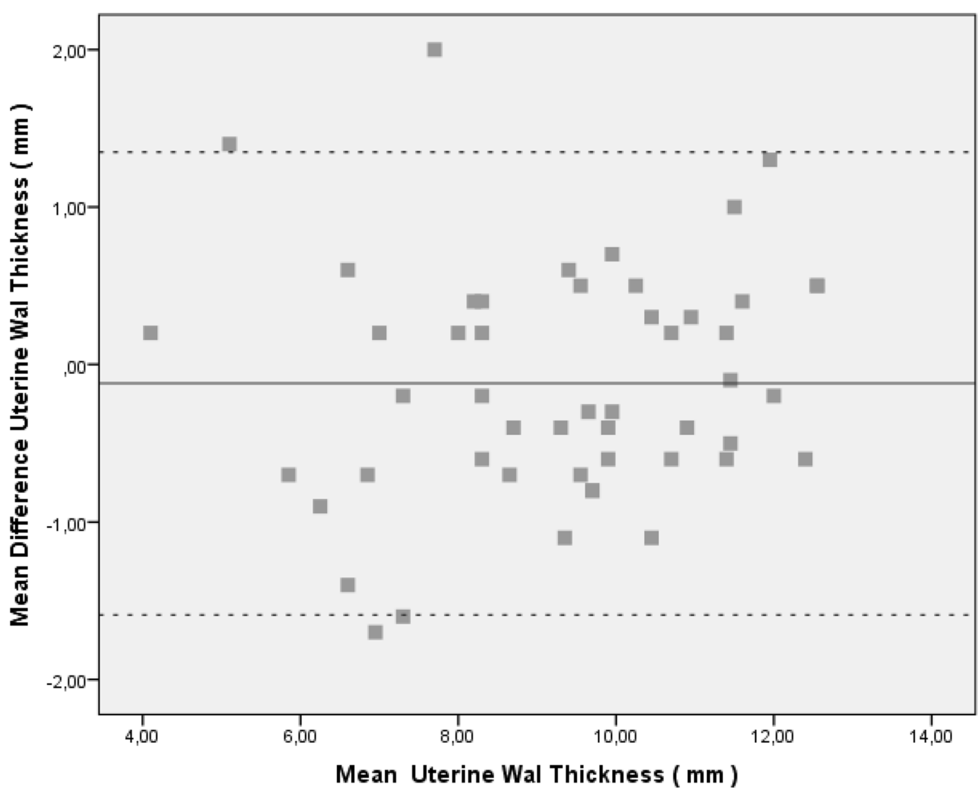

Fig. 1 Interobserver variability of uterine wall thickness (mm). Mean difference: − 0.12 mm (95% CI − 1.59 to 1.35)

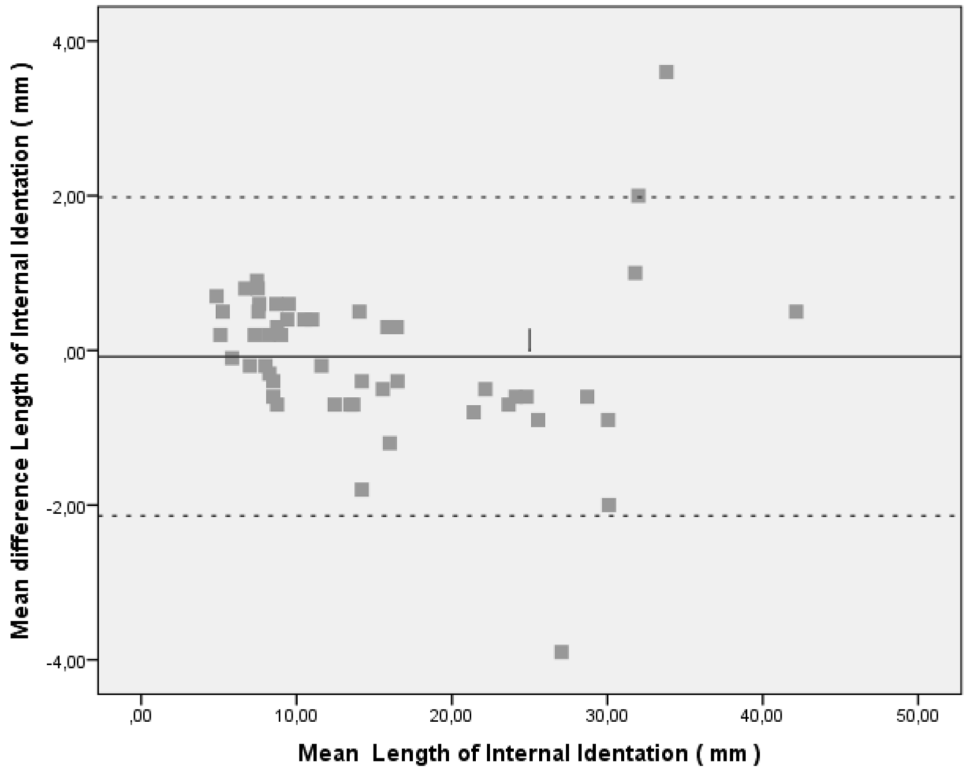

Fig. 2 Interobserver variability of length of internal indentation (mm). Mean difference: −0.08 mm (95% CI −2.14 to 1.98)

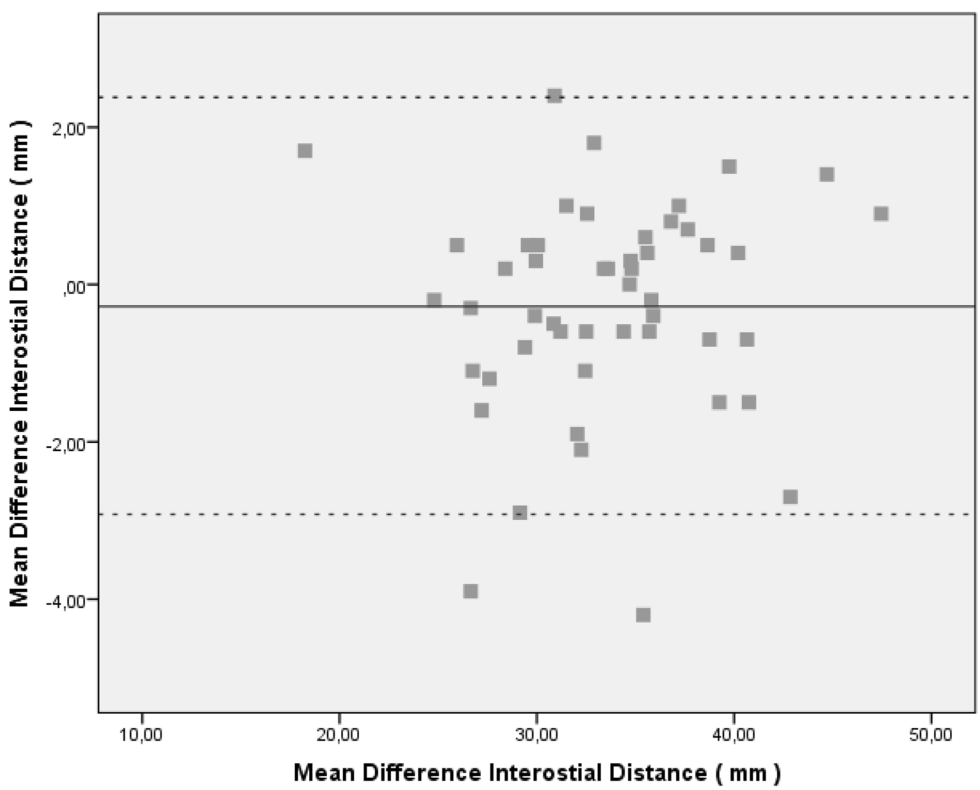

Fig. 3 Interobserver variability of interostial distance (mm). Mean difference: −0.27 mm (95% CI −2.92 to 2.38)

Table 7 Concordance between 3D US and HL

3D ultrasound	Hysteroscopy/laparoscopy						
	U0	U1	U2	U3	U4	U5	Total
U0							
N	1	0	0	0	0	0	1
%	1.6	0.0	0.0	0.0	0.0	0.0	1.6
U1							
N	0	3	0	0	0	0	3
%	0.0	4.8	0.0	0.0	0.0	0.0	4.8
U2							
N	1	0	49	0	0	0	50
%	1.6	0.0	78.9	0.0	0.0	0.0	80.5
U3							
N	0	0	0	3	0	0	3
%	0.0	0.0	0.0	4.8	0.0	0.0	4.8
U4							
N	0	0	0	0	4	0	4
%	0.0	0.0	0.0	0.0	6.5	0.0	6.5
U5							
N	0	0	0	0	0	1	1
%	0.0	0.0	0.0	0.0	0.0	1.6	1.6
Total							
N	2	3	49	3	4	1	62
%	3.2	4.8	78.9	4.8	6.5	1.6	100.0

U0, normal uterus; U1, dysmorphic uterus; U2, septate uterus; U3, bicorporeal uterus; U4, hemi-uterus; U5, aplastic uterus

hemi-uterus (Class U4) in 4/62 (6.5%) patients, and aplastic uterus (Class U5) in 1/62 (1.6%) patient.

Agreement between 3D US and H & L

Exact concordance (Kappa index 0.945) between 3D US and H & L about the type and classification of uterine anomaly was verified in 61/62 cases (Table 7). Only in one case there was discordance between 3D US and endoscopy: a case with normal uterus was diagnosed with the 3D US as a partial septum.

For the diagnosis of septate uteri, which was the most common anomaly, the sensitivity of 3D US was 100%, the specificity was 92.3%, the PPV was 98% and the NPV was 100%, with kappa index 0.950. For bicorporeal, dysmorphic uterus, hemi-uteri or unicorporeal and aplastic uterus the

Table 8 Diagnostic accuracy of 3D US compared to the hysteroscopy/laparoscopy in the diagnosis of uterine anomalies

	Sensitivity	Specificity	PPV	NPV	Kappa value
U0	0.500 (1/2) (0.03–0.97)	1.00 (60/60) (0.92–1.00)	1.00 (1/1) (0.05–1.00)	0.984 (60/61) (0.90–1.00)	0.659 (0.04–1.00)
U1	1.00 (3/3) (0.31–0.99)	1.00 (59/59) (0.92–1.00)	1.00 (3/3) (0.31–1.00)	1.00 (59/59) (0.92–1.00)	1.00 (1.00–1.00)
U2	1.00 (49/49) (0.91–1.00)	0.923 (12/13) (0.62–1.00)	0.980 (49/50) (0.88–1.00)	1.00 (12/12) (0.70–1.00)	0.950 (0.85–1.00)
U3	1.00 (3/3) (0.31–1.00)	1.00 (59/59) (0.92–1.00)	1.00 (3/3) (0.31–1.00)	1.00 (59/59) (0.92–1.00)	1.00 (1.00–1.00)
U4	1.00 (4/4) (0.40–1.00)	1.00 (58/58) (0.92–1.00)	1.00 (4/4) (0.40–1.00)	1.00 (58/58) (0.92–1.00)	1.00 (1.00–1.00)
U5	1.00 (1/1) (0.05–1.00)	1.00 (61/61) (0.92–1.00)	1.00 (1/1) (0.05–1.00)	1.00 (61/61) (0.92–1.00)	1.00 (1.00–1.00)

PPV positive predictive value, *NPV* negative predictive value, *U0* normal uterus, *U1* dysmorphic uterus, *U2* septate uterus, *U3* bicorporeal uterus, *U4* hemi-uterus, *U5* aplastic uterus sensitivity, specificity, PPV and NPV were all 100% with $K = 1.00$. Overall, 3D US showed perfect diagnostic accuracy (Kappa index = 0.945) in the detection of congenital uterine anomalies (Table 8).

Discussion

This study supports that 3D US is very accurate in the diagnosis and classification of congenital uterine malformations. The overall diagnostic accuracy of the method is very high (Kappa index 0.945 (0.839–1.000). Moreover, 3D US demonstrates an excellent reproducibility in the diagnostic approach of congenital uterine malformations (ICC > 0.934). This is a significant result that exhorts clinicians to increase both the use of 3D US for the investigation of congenital uterine diseases and the use of ESHRE/ESGE classification as a simplified and reliable instrument for reproducible and punctuate results.

Study results compared to the literature

This study used a highly selected group to compare 3D US with endoscopy. All our patients were initially selected and enrolled as potentially abnormal cases according to 2D US. Our findings indicate that 3D US is extremely accurate in the diagnosis of congenital uterine malformations, as it was endoscopically confirmed in 61 out of 62 abnormal cases (98.4%). The most frequent anomaly diagnosed was the septate uterus; 3D US showed 100% sensitivity, 92.3% specificity, and a Kappa index 0.950. Faivre et al. used a similarly highly selected population consisting of 20 patients with septate uterus and 11 patients with bicornuate uterus and used the ASRM classification for uterine anomalies. 3D US was extremely accurate in the diagnosis of uterine anomalies in that study, as it diagnosed correctly all the cases compared to operative endoscopy [17]. Ghi et al. had similar results in their study: the use of 3D US for the diagnosis and classification of congenital uterine anomalies in nulliparae women with three or more miscarriages showed very high accuracy when compared to endoscopy [16]. Graupera et al. compared 3D US to MRI performance using ESHRE/ESGE classification in 60 patients with known uterine malformations and found a very good agreement between the two methods in almost all the classes of congenital uterine anomalies [19].

3D US reproducibility

The results of this study also showed almost perfect agreement between the two observers in the diagnosis and classification of congenital uterine anomalies using 3D US. The reproducibility of measurements of the uterine wall thickness, the length of internal indentation and the interostial distance in the cases with septate uterus showed high levels of agreement between the two observers (ICC > 0.95). These results corroborate the relevant findings in the published literature. Salim et al. examined the reproducibility of 3D US measurements of interostial distance, cavity indentation and cavity length and showed that the intra- and interobserver variability were satisfactory ($k = 0.97$, 95% CI 0.94–1.0) [20]. Moreover, Saravelos et al. compared different 3D US uterine measurements and described intra- and interobserver variability for diagnosis of the most common congenital uterine anomalies. The measurements of the fundal wall, the cavity indentation, the interostial distance, the cavity angle and percentage of indentation showed high levels of intra- and interobserver reproducibility with intra-CC and inter-CC of 0.95 or over. On the other hand, measurements of the lateral uterine walls had lower reproducibility and the authors commented about the appropriateness of the 3D US in the evaluation of the uterine wall thickness [21]. Ludwin et al. reported a high level of inter-rater reliability during their study, in which they evaluated the reliability of the 3D ultrasonographic measurements of the uterine wall thickness, of the internal fundal indentation and of the external cleft. They included 50 patients with uterine anomalies and 62 women with a normal uterus and they used the ESHRE-ESGE system for the classification of uterine anomalies. The inter-rater reliability was > 0.99 (0.991; 95% CI 0.987–0.994) for the internal fundal indentation, and 0.95–0.99 (0.968; 95% CI 0.915–0.988 and 0.958; 95% CI 0.939–0.971 for the external cleft and the uterine wall thickness, respectively [22].

Cases of discordance: explanations, comparisons to literature

The 3D US diagnosis proved to be inaccurate only in one of our cases. This refers to a case of normal uterus, which was diagnosed by 3D US as a partial septate uterus. Although not reported in the literature, it is possible that a fundal myometrial contraction may temporarily increase the thickness and the outline of the fundal myometrium and may cause a temporary and usually mild indentation of the uterine cavity. In case of doubt, a repeat scan may be useful [23]. So, we speculate that false-positive finding in our study was due to our limited experience at the beginning of the study: a temporary contraction of fundal myometrium caused a mild indentation of the uterine cavity and, as a result, the thickness of fundal myometrium was evaluated as partial uterine septum; the reassessment of the same case at hysteroscopy led to the correct diagnosis of a normal uterus.

In the literature, Bermejo et al. suggested that 3D US performs suboptimally in the study of the lower part of the uterus and the cervix. In such cases, correct diagnosis is set only after confirmation with clinical inspection of the cervix. On the other hand, MRI can discriminate a septum in the lower part of the uterus and, in the cervix, therefore might be more suitable method for these anomalies [24]. Our study included seven patients with significant cervical abnormalities: in all these cases, 3D US confirmed the diagnosis. Therefore, our findings support that 3D US could be used in the diagnostic approach of congenital cervical anomalies as well.

ESHRE/ESGE classification: advantages and disadvantages

The methodology of the current study included the ESHRE/ESGE classification for uterine anomalies as the tool for the description of the abnormal findings. The general characteristics of this classification are that (a) the basis for the categorization is anatomy, (b) the basis for the main classes is deviations from same embryological origin, (c) the basis for the main sub-classes is variations expressing different extent of uterine deformity, and (d) cervical and vaginal anomalies are described independently [9]. The comprehensiveness of this classification was demonstrated in the precise categorization of 38 out of 39 types of unclassified according to the ASRM system congenital malformations published in the literature [25]. The results from our study show that the ESHRE/ESGE classification provides effective objective parameters for categorization of female congenital anomalies and allows us to classify the uterine malformations to classes and sub-classes according to the female genital tract anatomy. Moreover, in a recent retrospective study it was described that the ESHRE/ESGE classification covered the majority of the uterine and vaginal anomalies whereas the ASRM classification performed suboptimal in the description of cervical and vaginal anomalies [28].

However, Ludwin et al. commented that the ESHRE/ESGE system allows a relative overdiagnosis of septate uterus when compared to the ASRM classification; the authors considered that the use of the internal fundal indentation and the ratio of the internal fundal indentation to endometrial thickness in the ESHRE/ESGE system caused an increased false-positive rate of septate uterus compared to the more precise in this category ASRM classification. This might be quite important as any overdiagnosis of septate uterus may lead to unnecessary overtreatment [26]. Similar conclusions have been reported by Somayya et al. [27]. Nevertheless, the results of these studies should be cautiously interpreted; as in both studies, no gold standard method was used for the final comparison in between the ESHRE/ESGE and the ASRM classification. More importantly, it appears that the former study was highly biased by the selection criteria used for the included patients, where many normal patients were studied and this had as a consequence profoundly lower number of patients with internal indentation around the cut-off values between normal vs arcuate and normal vs septate uterus. Interestingly, the same group in the same study found high reproducibility of the parameters used for the estimation of the uterine anatomy using 3D US [22]. Based on these findings and on the findings of Saravelos et al. [21], it appears that the use of uterine wall thickness as a reference parameter when performing 3D US could be used as a reliable tool for objective estimation of the uterine deformities.

Strengths and limitations of the study

The main strengths of the study are (a) the prospective design, (b) that 3D US diagnosis of uterine anomalies is compared to hysteroscopy with laparoscopy that is considered as the gold standard in the diagnosis of uterine malformations, and (c) the use of ESHRE/ESGE classification which allows the employment of highly reproducible sonographic parameters.

There are some limitations of this study. First, it is the low statistical power, especially because of the small number of patients with other uterine anomalies, excluding the septate uterus. However, the prevalence of these uterine malformations is anyway low. Second, the study included a group of patients with suspected congenital uterine anomaly after 2D US examination, not the general population.

Conclusion

In conclusion, according to our data, transvaginal 3D ultrasonography is an accurate method for the diagnosis and classification of congenital uterine anomalies. Further studies, where prospective comparison of the use of 3D US and other currently available diagnostic modalities in the diagnosis of uterine malformations, will help in clarifying the exact role of each method in the investigation of congenital anomalies. The ESHRE/ESGE classification of uterine malformations appears to be a useful tool in the diagnostic approach of these conditions.

Authors' contribution AK: Data collection, Manuscript writing, data analysis. TM: Manuscript writing—review and editing, protocol development. GFG: Protocol development, data collection, conceptualization, investigation, project administration, supervision. AK: Data collection, investigation. TDT: Data collection, investigation. AS: Protocol development, supervision. BCT: Supervision. APA: Protocol development, data collection, conceptualization, investigation, project administration, supervision, validation.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Grimbizis GF, Campo R (2010) Congenital malformations of the female genital tract: the need for a new classification system. Fertil Steril 94:401–407
2. Saravelos SH, Cocksedge KA, Li T-C (2008) Prevalence and diagnosis of congenital uterine anomalies in women with reproductive failure: a critical appraisal. Hum Reprod Update 14:415–419
3. Chan YY, Jayaprakasan K, Zamora J, Thornton JG, Raine-Fenning N, Coomarasamy A (2011) The prevalence of congenital uterine anomalies in unselected and high-risk populations: a systematic review. Hum Reprod Update 17:761–771
4. Buttram VC, Gibbons WE (1979) Mullerian anomalies: a proposed classification (an analysis of 144 cases). Fertil Steril 32:40–46
5. The American Fertility Society (1988) The AFS classification of adnexal adhesions, distal tubal occlusion, tubal occlusion secondary to tubal ligation, tubal pregnancies, Mullerian anomalies and intrauterine adhesions. Fertil Steril 49:944–955
6. Acien P, Acien M, Sanchez-Ferrer M (2004) Complex malformations of the female genital tract: new types and revision of classification. Hum Reprod 19:2377–2384
7. Oppelt P, Renner SP, Brucker S, Strissel PL, Strick R, Oppelt PG, Doerr HG, Schott GE, Hucke J, Wallwiener D, Beckmann MW (2005) The VCUAM (vagina cervix uterus adnex associated malformation) classification: a new classification for genital malformations. Fertil Steril 84:1493–1497
8. Grimbizis GF, Gordts G, Di Spiezio SA, Brucker S, De Angelis C, Gergolet M, Li T-C, Tanos V, Brölmann H, Gianaroli L, Campo R (2013) The ESHRE/ESGE consensus on the classification of female genital tract congenital malformations. Hum Reprod 28:2032–2044
9. Grimbizis GF, Gordts G, Di Spiezio SA, Brucker S, De Angelis C, Gergolet M, Li T-C, Tanos V, Brölmann H, Gianaroli L, Campo R (2013) The ESHRE/ESGE consensus on the classification of female genital tract congenital malformations. Gynecol Surg 10:199–212
10. Grimbizis GF, Di Spiezio Sardo A, Saravelos SH, Gordts S, Exacoustos C, Van Schoubroeck D, Bermejo C, Amso NN, Nargund G, Timmermann D, Athanasiadis A, Brucker S, De Angelis C, Gergolet M, Li TC, Tanos V, Tarlatzis B, Farquharson R, Gianaroli L, Campo R (2016) The Thessaloniki ESHRE/ESGE consensus on diagnosis of female genital anomalies. Gynecol Surg 13:1–16
11. Byrne J, Nussbaum-Blask A, Taylor WS, Rubin A, Hill M, O'Donnel R, Shulman S (2000) Prevalence of Mullerian duct anomalies detected at ultrasound. Am J Med Genet 94:9–12
12. Troiano RN, McCarthy SM (2004) Mullerian duct anomalies: imaging and clinical issues. Radiology 233:19–34
13. Nicolini U, Bellotti M, Bonazzi B, Zamberletti D, Candiani GB (1987) Can ultrasound be used to screen uterine malformations? Fertil Steril 47:89–93
14. Salim R, Woelfer B, Backos M, Regan L, Jurkovic D, Letterie GS, Haggerty M, Lindee G (1995) A comparison of pelvic ultrasound and magnetic resonance imaging as diagnostic studies for mullerian tract abnormalities. Int J Fertil Menopausal Stud 40:34–38
15. Jurkovic D, Geipel A, Gruboeck K, Jauniaux E, Natucci M, Campbell S (1995) Three-dimensional ultra- sound for the assessment of uterine anatomy and detection of congenital anomalies: a comparison with hysterosalpingography and two-dimensional sonography. Ultrasound Obstet Gynecol 5:233–237
16. Ghi T, Casadio P, Kuleva M, Perrone AM, Savelli L, Giunchi S, Meriggiola MC, Gubbini G, Pilu G, Pelusi C, Pelusi G (2009) Accuracy of three-dimensional ultra- sound in diagnosis and classification of congenital uterine anomalies. Fertil Steril 92:808–813
17. Faivre E, Fernandez H, Deffieux X, Gervaise A, Frydman R, Levaillant JM (2012) Accuracy of three- dimensional ultrasonography in differential diagnosis of septate and bicornuate uterus compared with office hysteroscopy and pelvic magnetic resonance imaging. J Minim Invasive Gynecol 19(1):101–106
18. Ludwin A, Pitynski K, Ludwin I, Banas T, Knafel A (2013) Two- and three- dimensional ultrasonography and sonohysterography versus hysteroscopy with laparoscopy in the differential diagnosis of septate, bicornuate, and arcuate uteri. J Minim Invasive Gynecol 20(1):90–99
19. Graupera B, Pascual MA, Hereter L, Browne JL, Úbeda B, Rodriguez I, Pedrero C (2015) Accuracy of three-dimensional ultrasound compared with magnetic resonance imaging in diagnosis of Mullerian duct anomalies using ESHRE-ESGE consensus on the classification of congenital anomalies of the female genital tract. Ultrasound Obstet Gynecol 46(5):616–622
20. Salim R, Woelfer B, Backos M, Regan L, Jurkovic D (2003) Reproducibility of three-dimensional ultrasound diagnosis of congenital uterine anomalies. Ultrasound Obstet Gynecol 21(6):578–582
21. Saravelos SH, Li TC (2015) Intra- and inter-observer variability of uterine measurements with three-dimensional ultrasound and implications for clinical practice. Reprod Biomed Online 31(4):557–564
22. Ludwin A, Ludwin I, Kudla M, Kottner J (2015) Reliability of the European Society of Human Reproduction and Embryology/European Society for Gynaecological Endoscopy and American Society for Reproductive Medicine classification systems for congenital uterine anomalies detected using three-dimensional ultrasonography. Fertil Steril 104:688–697
23. Van Schoubroeck D, Van den Bosch T, Timmerman D (2015) 3D ultrasound. In: Grimbizis GF, Campo R, Tarlatzis BC, Gordts S (eds) Female genital tract congenital malformations. Classification, diagnosis and management. Springer, Berlin, pp 79–97
24. Bermejo C, Ten MP, Cantarero R, Diaz D, Pedregosa JP, Barrón E, Labrador E, López LR (2010) Three-dimensional ultrasound in the diagnosis of Mullerian duct anomalies and concordance with magnetic resonance imaging. Ultrasound Obstet Gynecol 35(5):593–601
25. Di Spiezio Sardo A, Campo R, Gordts S, Spinelli M, Cosimato C, Tanos V, Brucker S, Li TC, Gergolet M, De Angelis C, Gianaroli L, Grimbizis G (2015) The comprehensiveness of the ESHRE/ESGE classification of female genital tract congenital anomalies: a systematic review of cases not classified by the AFS system. Hum Reprod 30(5):1046–1058
26. Ludwin A, Ludwin I (2015) Comparison of the ESHRE-ESGE and ASRM classifications of Mullerian duct anomalies in everyday practice. Hum Reprod 30(3):569–580
27. Somayya M, Reda A, Atia H (2016) Performance of the ESHRE/ESGE classification in differentiating anomalies of double uterine cavity in comparison with the ASRM classification. Middle East Fertil Soc J 21(2):75–81
28. Heinonen PK (2016) Distribution of female genital tract anomalies in two classifications. Eur J Obstet Gynecol Reprod Biol 206:141–146