Thieme

Can Classifications Adequately Represent Genital Malformations?

EVA Study (ESHRE/ESGE | VCUAM | AFS) – A Prospective Multicenter Study to Evaluate the Current Female Genital Malformation Classifications

Können Klassifikationssysteme genitale Fehlbildungen adäquat darstellen?

EVA Study (ESHRE/ESGE | VCUAM | AFS) – eine prospektive multizentrische Studie zur Evaluierung aktueller Klassifikationssysteme für weibliche genitale Fehlbildungen

Authors

Stephanie Kiblboeck[1], Peter Oppelt[1], Patricia Oppelt[2], Raimund Stein[3], Stefanie Ommer[4], Roman Pavlik[5], Katharina Rall[6], Kuralay Kongrtay[7], Helga Wagner[8,9], Philipp Hermann[9], Philip Sebastian Trautner[1]

Affiliations

1 Department of Gynecology, Obstetrics and Gynecological Endocrinology, Kepler University Hospital Linz, Johannes Kepler Universität Linz, Linz, Austria
2 Department of Gynecology, Erlangen University Hospital, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
3 Department of Pediatric, Adolescent and Reconstructive Urology, Medical Faculty Mannheim, University Medical Center Mannheim, Heidelberg University, Mannheim, Germany
4 Department of Pediatric Surgery, University Hospital Jena, Jena, Germany
5 TFP Fertility Wels, Thalheim bei Wels, Austria
6 Department of Gynecology and Obstetrics, University Hospital Tübingen, Tübingen, Germany
7 Clinical Academic Department of Women's Health, Corporate fund "University Medical Center", Astana, Kazakhstan
8 Department of Applied Statistics, Johannes Kepler University Linz, Linz, Austria
9 Center for Clinical Studies (CCS Linz), Johannes Kepler University Linz, Linz, Austria

Key words

EVA study, female genital malformation, uterine malformation, genital classification

Schlüsselwörter

EVA-Studie, weibliche genitale Fehlbildungen, Gebärmutterfehlbildung, genitale Klassifikation

received 10.11.2022
accepted after revision 20.2.2023
published online 29.3.2023

Bibliography

Geburtsh Frauenheilk 2023; 83: 827–834
DOI 10.1055/a-2043-9982
ISSN 0016-5751

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence

Stephanie Kiblboeck, M.D.
Department of Gynecology, Obstetrics and Gynecological Endocrinology
Kepler University Hospital Linz, Johannes Kepler Universität Linz
Altenberger Strasse 69
4040 Linz, Austria
stephanie.kiblboeck@gmx.at

Additional material is available at https://doi.org/10.1055/a-2043-9982.

ABSTRACT

Introduction Genital malformations are a common clinical occurrence that can be represented using different classifications. Reproducibility is an essential quality characteristic for a classification, and it plays an important role, especially in consultations and the treatment of infertile patients and in obstetric management. The aim of this study is to demonstrate the reproducibility and clinical practicality of three commonly used classifications: the ESHRE/ESGE (European Society of Human Reproduction and Embryology/ European Society for Gynecological Endoscopy), VCUAM (Vagina Cervix Uterus

Adnex-associated Malformation), and AFS (American Fertility Society) classifications.

Materials and Methods Sixty-five patients with female genital malformations were included in this prospective, multicenter, exploratory, observational study. All participants underwent a clinical examination and a medical interview. The investigators were instructed to classify the presenting malformations according to the ESHRE/ESGE, VCUAM, and AFS classifications using a structured questionnaire. Investigators were asked whether the malformation could be reproducibly classified (yes/no) and about the grade (grade 1–5 from "very good" to "deficient") they would assign to each classification. Classification assessment was queried for vagina, cervix, uterus, adnexa, and associated malformations and was scored from 1 to 5.

Results Reproducibility was rated as 80 % (n = 52/65), 92.3 % (n = 60/65), and 56.9 % (n = 37/65) for the ESHRE/ESGE, VCUAM, and AFS classification, respectively. ESHRE/ESGE, VCUAM and AFS were rated as "very good" or "good" for 83.3 %, 89.2 %, and 10.8 % of vaginal malformations; for 75.8 %, 87.5 %, and 24.2 % of cervical malformations; and for 89.7 %, 89.5 %, and 86.2 % of uterine malformations, respectively. VCUAM was rated as "very good" or "good" for 77.8 % and 69.6 % of adnexal malformations and associated malformations, respectively. ESHRE/ESGE and AFS were rated as "sufficient" or "deficient" for 100 % and 75 % of adnexal malformations and for 77.3 % and 69.6 % of associated malformations, respectively.

Conclusion The prospective multicenter EVA (ESHRE/ESGE | VCUAM | AFS) study revealed that the organ-based ESHRE/ESGE and VCUAM classifications of female genital malformations perform better in terms of reproducibility as well as in the assessment of individual compartments than the non-organ-based AFS classification.

ZUSAMMENFASSUNG

Einleitung Genitale Fehlbildungen treten häufig in der klinischen Praxis auf. Sie werden mithilfe verschiedener Klassifikationssysteme beschrieben. Die Reproduzierbarkeit sollte ein wesentliches Qualitätsmerkmal für jede Klassifikation sein, da die korrekte Klassifizierung eine wichtige Rolle besonders in der Beratung und Behandlung infertiler Patientinnen und beim geburtshilflichen Management spielt. Ziel dieser Studie war es, die Reproduzierbarkeit und klinische Praktikabilität von 3 häufig verwendeten Klassifikationssystemen zu überprüfen: das ESHRE/ESGE- (European Society of Human Reproduction and Embryology/ European Society for Gynecological Endoscopy), das VCUAM- (Vagina Cervix Uterus Adnex-associated Malformation) und das AFS-Klassifikationssystem (AFS: American Fertility Society).

Material und Methoden Es wurden 65 Patientinnen mit genitalen Fehlbildungen in diese prospektive multizentrische Sondierungs- und Beobachtungsstudie aufgenommen. Alle Teilnehmerinnen wurden klinisch untersucht und hatten ein medizinisches Gespräch. Die Untersucher*innen wurden angewiesen, die vorliegenden Fehlbildungen gemäß den ESHRE/ESGE-, VCUAM- und AFS-Klassifikationen mithilfe eines strukturierten Fragebogens einzustufen. Die Untersucher*innen wurden gefragt, ob die jeweilige Fehlbildung reproduzierbar klassifiziert werden konnte (ja/nein) und sollten die Qualität (Note 1–5 von „Sehr gut" bis „Nicht genügend") der jeweiligen Klassifikation beurteilen. Diese Bewertung wurde für Vagina, Zervix, Uterus, Adnexe sowie assoziierte Fehlbildungen durchgeführt.

Ergebnisse Die Reproduzierbarkeit wurde mit 80 % (n = 52/65), 92,3 % (n = 60/65) resp. 56,9 % (n = 37/65) für die ESHRE/ESGE-, VCUAM- bzw. AFS-Klassifikation bewertet. Die ESHRE/ESGE-, VCUAM- und AFS-Klassifikationen wurden jeweils als „Sehr gut" oder „Gut" bei 83,3 %, 89,2 % resp. 10,8 % der vaginalen Fehlbildungen, bei 75,8 %, 87,5 % resp. 24,2 % der zervikalen Fehlbildungen und bei 89,7 %, 89,5 % resp. 86,2 % der uterinen Fehlbildungen bewertet. VCUAM wurde als „Sehr gut" oder „Gut" bei 77,8 % resp. 69,6 % der Fehlbildungen der Adnexe und der assoziierten Fehlbildungen eingestuft. ESHRE/ESGE und AFS wurden als „Genügend" oder „Nicht genügend" bei 100 % bzw. 75 % der Fehlbildungen der Adnexe und bei 77,3 % bzw. 69,6 % der assoziierten Fehlbildungen eingestuft.

Schlussfolgerung Die prospektive multizentrische EVA-Studie (EVA: ESHRE/ESGE | VCUAM | AFS) zeigt, dass die organbezogenen ESHRE/ESGE- und VCUAM-Klassifikationen weiblicher genitaler Fehlbildungen im Hinblick auf die Reproduzierbarkeit sowie die Evaluierung individueller Kompartimente besser abschneiden als die nicht organbezogene AFS-Klassifikation.

Abbreviations

AFS	American Fertility Society
ASRM	American Society for Reproductive Medicine
ESGE	European Society for Gynecological Endoscopy
ESHRE	European Society of Human Reproduction and Embryology
FIGO	Fédération Internationale de Gynécologie et d'Obstétrique
Grav	Gravidity
IQR	Interquartile Range
TNM	TNM (Tumor, nearby Lymph Nodes, Metastasis) Classification of Malignant Tumors
VCUAM	Vagina Cervix Uterus Adnex–associated Malformation

Introduction

Female genital malformations are anomalies of the female genital tract, and their true incidence in the general population is difficult to ascertain. For this reason, a large incidence range from 0.2 to 6.7 % has been reported in the literature [1, 2, 3, 4]. In patients with infertility and recurrent miscarriage, the incidence of genital malformations is estimated to be even higher at 7% and 17%, respectively [1, 4, 5, 6].

Many of these malformations result from a developmental disorder of the Müllerian ducts, resulting in anomalies of the uterus, cervix, fallopian tubes, and proximal vagina [7]. Complex genital malformations may represent a combination of alterations of the Müllerian and Wolffian ducts and the urogenital sinus. Female genital malformations involving the distal vagina such as transverse vaginal septa or hymenal atresia are not due to anomalies of the Müllerian ducts [7, 8, 9].

Classifications

Buttram et al. (1979) divided Müllerian duct anomalies into six groups, focusing on uterine malformations with similar clinical manifestations, treatment, and pregnancy outcomes. This classification was revised by the AFS in 1988, with the main difference being that the arcuate uterus was delineated as a separate subgroup from the subseptate uterus. The malformations within each of the seven groups are presented with schematic illustrations and may be supplemented by a subgroup. Additional findings such as malformations of the vagina, cervix, tubes, and kidneys, are presented descriptively [10, 11]. Acien (1992) succeeded in explaining the pathogenesis of most anomalies of the female genital tract using an embryological hypothesis, and thus a classification was subsequently developed by Acien et al. (2004) to categorize Müllerian anomalies according to their embryological origin [7, 12]. In 2011, this classification was updated to include six groups

1. agenesis or hypoplasia of an entire urogenital ridge,
2. mesonephric anomalies,
3. isolated Müllerian anomalies,
4. gubernaculum dysfunctions,
5. anomalies of the urogenital sinus, and
6. combinations of malformations [13].

In their recent editorial, Acien et al. (2022) point out that knowledge of the correct genitourinary embryology is essential for the diagnosis, understanding, and subsequent management of genital malformations, particularly complex malformations that lead to gynecologic and reproductive disorders [14].

Oppelt et al. (2005) established the VCUAM classification, which is based on external and internal female genital malformations, including malformations that are not due to Müllerian duct anomalies. This classification divides malformations into five main groups based on anatomy: vagina (V), cervix (C), uterus (U), adnexa (A), and associated malformations (M). The more pronounced the malformation, the higher the numerical classification. The highest number in each group indicates atresia or aplasia [15].

Grimbizis et al. (2013) published the ESHRE/ESGE consensus on the classification of congenital anomalies of the female genital tract, based on anatomical changes. Uterine malformations of the same embryological origin are divided into the main groups, and different alterations are divided into subgroups according to their severity. Cervical and vaginal malformations can be classified into independent supplementary subclasses. Associated anomalies of non-Müllerian origin are presented descriptively [16].

To sum up, the ESHRE/ESGE and VCUAM classifications are organ-based, which means that individual compartments can be independently classified, comparable to the TNM classification of malignant tumors [15]. The AFS classification provides an overall picture of the malformation, most comparable to the FIGO (Fédération Internationale de Gynécologie et d'Obstétrique) classification in oncology.

Reproducibility is an essential quality characteristic for classification, and it plays an important role, especially in consultations and the treatment of patients with infertility and in obstetric management. To date, no prospective studies have examined or compared the reproducibility and clinical practicality of these three commonly used classifications. Therefore, we conducted the EVA (ESHRE/ESGE | VCUAM | AFS) study, a prospective, international, multicenter study which included an opinion survey.

Material and Methods

Between December 2019 and February 2022, 65 female patients (median age: 22.37 years [interquartile range IQR: 12.02]) with genital malformations were enrolled in six international participating experienced tertiary care units for patients with genital malformations as part of a prospective exploratory study (the EVA study) which included an opinion survey. Most physicians (n = 12) were gynecologists (83.3 %, n = 10). A pediatric surgeon (8.3 %) and a pediatric urologist (8.3 %) also participated in the study. Thirty-two patients were assessed in center 1 by six investigators, mainly by investigators 1 and 2. Center 2 evaluated four, center 3 eleven, center 4 three, center 5 eleven and center 6 four patients, respectively, which represents a skewed distribution of the study participants with respect to study centers and investigators. The details of the patients' baseline characteristics and the distribution of investigators are shown in ▶ **Table 1** and ▶ **Table 2**. The local ethics committee of each center approved this study and informed consent was obtained from all patients or their guardians.

EVA study questionnaire and assessment

In the EVA study, a questionnaire was designed to investigate the subjective reproducibility and clinical practicality of each classification. All participants underwent a clinical examination and a medical interview. In addition to the general characteristics of the patient, such as age, symptoms, occurrence of symptoms, and medical history, the patient's condition following genital malformation correction and obstetric history were also collected. (The corresponding questionnaire is attached to the supplemental data.)

▶**Table 1** Baseline characteristics of patients and physicians. Median and IQR are provided for metric variables (with more than three observations per group, otherwise NA). Absolute and relative frequencies in percent are computed for nominal and ordinal variables for the total study population as well as per study center. Categories larger than 2 are merged for parity, gravidity, and number of miscarriages. Note that only observations with valid observations are used to calculate percentage values. P values of Kruskal-Wallis test for differences in metric data between study centers and of Fisher's exact test for nominal or grouped ordinal data are provided.

	Total	Center 1	Center 2	Center 3	Center 4	Center 5	Center 6	P value
Patients, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
Age, median (IQR)	22.37 (12.02)	22.99 (9.92)	NA (NA)	10.17 (21.47)	15.1 (2.94)	NA (NA)	NA (NA)	< 0.001
Gravidity, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
Grav 0, n (%)	52 (80)	25 (78.12)	4 (100)	11 (100)	2 (66.67)	7 (63.64)	3 (75)	0.487
Grav I, n (%)	6 (9.23)	3 (9.38)	0 (0)	0 (0)	1 (33.33)	2 (18.18)	0 (0)	
Grav ≥ II, n (%)	7 (10.77)	4 (12.5)	0 (0)	0 (0)	0 (0)	2 (18.18)	1 (25)	
Parity, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
Parity 0, n (%)	58 (89.23)	28 (87.5)	4 (100)	11 (100)	3 (100)	8 (72.73)	4 (100)	0.850
Parity I, n (%)	4 (6.15)	2 (6.25)	0 (0)	0 (0)	0 (0)	2 (18.18)	0 (0)	
Parity ≥ II, n (%)	3 (4.62)	2 (6.25)	0 (0)	0 (0)	0 (0)	1 (9.09)	0 (0)	
Miscarriage, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
Miscarriage 0, n (%)	57 (87.69)	27 (84.38)	4 (100)	11 (100)	2 (66.67)	9 (81.82)	4 (100)	0.576
Miscarriage I, n (%)	5 (7.69)	2 (6.25)	0 (0)	0 (0)	1 (33.33)	2 (18.18)	0 (0)	
Miscarriage ≥ II, n (%)	3 (4.62)	3 (9.38)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Familial malformation, n	**64**	**31**	**4**	**11**	**3**	**11**	**4**	
No familial malformation, n (%)	59 (92.19)	27 (87.1)	4 (100)	11 (100)	3 (100)	10 (90.91)	4 (100)	0.929
Familial malformation, n (%)	5 (7.81)	4 (12.9)	0 (0)	0 (0)	0 (0)	1 (9.09)	0 (0)	
Correction, n	**54**	**23**	**2**	**11**	**3**	**11**	**4**	
No correction, n (%)	34 (62.96)	19 (82.61)	1 (50)	0 (0)	3 (100)	7 (63.64)	4 (100)	< 0.001
Correction, n (%)	20 (37.04)	4 (17.39)	1 (50)	11 (100)	0 (0)	4 (36.36)	0 (0)	
Symptom, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
No symptoms, n (%)	18 (27.69)	11 (34.38)	0 (0)	3 (27.27)	0 (0)	4 (36.36)	0 (0)	0.574
Symptoms, n (%)	47 (72.31)	21 (65.62)	4 (100)	8 (72.73)	3 (100)	7 (63.64)	4 (100)	
Time in months, median (IQR)	12 (31)	26 (13.5)	30.5 (4.79)	24 (54.5)	0 (0.75)	6 (6)	NA (NA)	0.036
Syndrome, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
No syndrome, n (%)	53 (81.54)	31 (96.88)	0 (0)	7 (63.64)	3 (100)	9 (81.82)	3 (75)	< 0.001
Syndrome, n (%)	12 (18.46)	1 (3.12)	4 (100)	4 (36.36)	0 (0)	2 (18.18)	1 (25)	
Specialty, n	**65**	**32**	**4**	**11**	**3**	**11**	**4**	
Gynecology, n (%)	50 (76.92)	32 (100)	0 (0)	0 (0)	3 (100)	11 (100)	4 (100)	< 0.001
Pediatric surgery, n (%)	11 (16.92)	0 (0)	0 (0)	11 (100)	0 (0)	0 (0)	0 (0)	
Pediatric urology, n (%)	4 (6.15)	0 (0)	4 (100)	0 (0)	0 (0)	0 (0)	0 (0)	

The investigators were instructed to classify the presenting malformation using three classifications (ESHRE/ESGE, VCUAM, and AFS) and to rate the adequacy of the classification for assessing this malformation. Classification assessment was queried for vagina, cervix, uterus, adnexa, and associated malformations, and was scored from 1 to 5. A grade of 1 represented a very good assessment, whereas a grade of 5 represented an inadequate assessment. A grade of 6 indicated no malformation in the corresponding organ.

In addition, respondents were asked whether the malformation could be reproducibly classified, and which grade (grade 1–5) they would assign for the classification. Finally, a ranking of the clinical practicality of the three classifications was established. Rank 1 represented the most clinically practical classification, whereas rank 3 represented the least clinically practical classification.

Statistical analysis

The distributions of nominal and ordinal variables are presented as absolute and relative frequencies. Absolute and relative frequencies for grading malformations were computed based on the total number of present malformations per classification. Median and interquartile range (IQR) were provided for metric data, such as age. These summary statistics were computed overall, as well as per study center. Fisher's exact test was used to test the associations between nominal or grouped ordinal variables and the study center. Kruskal-Wallis test was used to compare the distribution of non-normal metric variables across the study centers. Stacked bar charts show the overall grade and the distribution of the assessment of present malformations (vagina, cervix, uterus, adnexa, and associated malformations) per classification method. The distribution of rankings for the overall classification methods as well as per study center is presented in the supplemental material using stacked bar charts of relative frequencies.

Results

Our study revealed that the subjective reproducibility determined by the investigators for all patients (n = 65) was higher for the organ-based classifications VCUAM (n = 60, 92.3 %) and ESHRE/ESGE (n = 52, 80 %) than for the non-organ-based AFS classification (n = 37, 56.9 %).

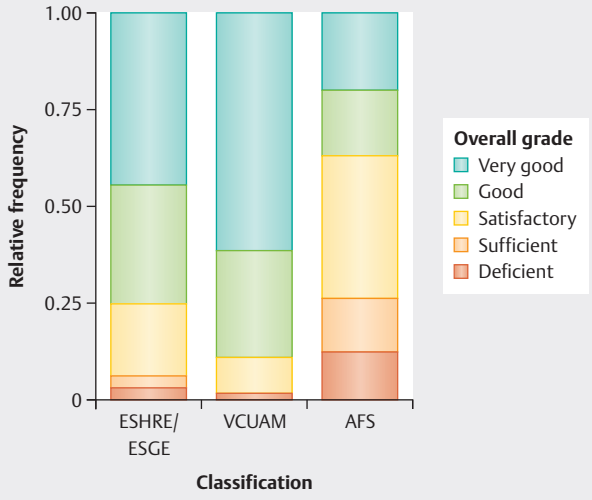
In addition, organ-based classifications in which malformations of each compartment could be described in detail performed better than the non-organ-based classification (▶ **Fig. 1**).

The assessment of individual compartments was considered in detail (▶ **Fig. 2**).

Vagina

The ESHRE/ESGE and VCUAM classifications were rated as "very good" or "good" for vaginal malformations in 83.3 % (n = 30/36 patients with vaginal malformations) and 89.2 % (n = 33/37) of patients, respectively. The AFS classification received this rating for 10.8 % of patients (n = 4/37).

The investigators noted that a detailed description of vaginal changes, including hymen changes, vaginal septa including length and extension (horizontal or transverse), obstruction, hypoplasia, atresia, and aplasia, with specification of the localization (left/right), and malformations of the urogenital sinus should be provided.



▶ **Fig. 1** Relative frequencies of overall grades per classification method.

Cervix

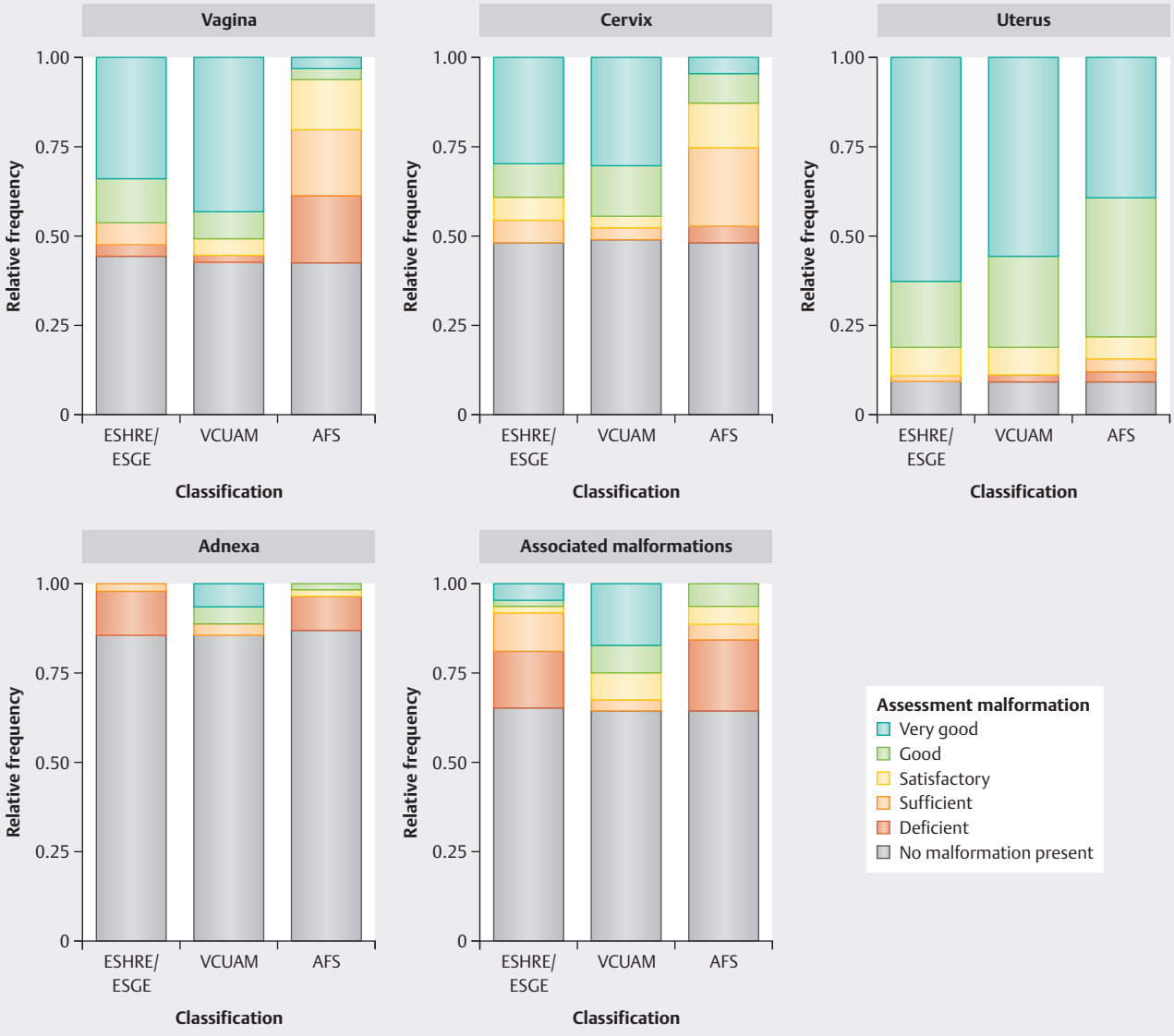
The ESHRE/ESGE and VCUAM classifications were rated as "very good" or "good" for cervical malformations in 75.8 % (n = 25/33) and 87.5 % (n = 28/32) of patients, respectively. The AFS classification received this rating for 24.2 % of the presented malformations (n = 8/33). The ability to differentiate between septate cervix and cervix duplex has been noted as a general criticism. In cases of atresia or aplasia, the specification of the localization (left/right) would be important for reproducibility.

Uterus

Uterine malformations could be described very accurately with all three classifications, and were rated as "very good" or "good" for ESHRE/ESGE in 89.7 % (n = 52/58 patients with uterine malformation), for VCUAM in 89.5 % (n = 51/57), and for AFS in 86.2 % (n = 50/58) of patients. The examiners noted that it was important that the classification included arcuate uterus and septate uterus, including length and extension of the septum. In cases with rudimentary horns, information whether they are communicating or non-communicating and specification of the localization (left/right) must be provided. Hypoplasia and aplasia should also be considered. An angle measurement was requested by one examiner to differentiate between arcuate and subseptate or septate and bicornuate uteri.

▶ **Table 2** Distribution of number of patients assessed per investigator.

													Total n (%)
Investigator	1	2	3	4	5	6	7	8	9	10	11	12	12 (100)
Patients assessed per investigator, n (%)	9 (13.9)	19 (29.2)	1 (1,5)	1 (1,5)	1 (1,5)	1 (1,5)	4 (6.2)	11 (16,9)	3 (4.6)	11 (16.9)	3 (4.6)	1 (1,5)	65 (100)



▶ **Fig. 2** Relative frequencies of assessments of malformations of the vagina, cervix, uterus, adnexa, and associated malformations per classification method.

Adnexa

The assessment of adnexal malformations was rated as “very good” or “good” by investigators for VCUAM in 77.8 % (n = 7/9) of patients with adnexal malformation. Assessments of adnexal malformations with ESHRE/ESGE and AFS were rated as “sufficient” or "deficient” in 100 % (n = 9/9) and 75 % (n = 6/8) of patients, respectively. Physicians would prefer an accurate list of possible adnexal malformations, such as aplasia, hypoplasia, hematosalpinx, sactosalpinx, and cysts.

Associated malformations

The best assessment in terms of associated malformations was obtained for VCUAM which was rated as “very good” or “good” in 69.6 % of cases (n = 16/23 patients with associated malformations). ESHRE/ESGE and AFS were rated as “sufficient” or "deficient” in 77.3 % (n = 17/22) and 69.6 % (n = 16/23) of patients, respectively, when associated malformations were assessed. As with adnexal malformations, the investigators recommend including a detailed list of possible associated malformations with specification of the localization (left/right).

The ranking results are shown in the supplemental material (see Supplemental Material, **Fig. S1**).

In general, the investigators stated that it was important to note in the classification whether the malformation had already been corrected. For complex malformations, the possibility of multiple selections as well as a space for drawings and descriptive explanations would be helpful.

Discussion

The aim of this EVA study was to compare the subjective reproducibility and clinical practicability of the ESHRE/ESGE, VCUAM, and AFS classifications. Reproducibility is an essential quality characteristic for classification, and it plays an important role, especially in consultations and the treatment of patients with infertility and in obstetric management.

The results of the EVA study demonstrate that organ-based classifications, i.e., the ESHRE/ESGE and VCUAM classifications, had a higher reproducibility than the non-organ-based AFS classification (ESHRE/ESGE 80 % and VCUAM 92.3 % vs. AFS 56.9 %).

This is because organ-based classifications where each compartment is described separately are comparable to the TNM classification in oncology and result in a significantly higher number of possible combinations. The ESHRE/ESGE classification yielded 349 possibilities for the presence of at least one malformation, whereas the VCUAM classification yielded 56 699 possibilities. The seven groups of the AFS classification, however, are unable to represent a malformation with similarly high level of reproducibility. Particularly in the case of complex malformations resulting from combinations of malformations of the Müllerian and Wolffian ducts or of the urogenital sinus, it appears that reproducibility is better when each compartment can be described separately in detail.

The results show the strengths and limitations of each classification (see ▶ **Fig. 2**). This has been repeatedly discussed in the literature. Uterine malformations can be adequately assessed using all three classifications, whereas cervical and vaginal malformations are better assessed using the ESHRE/ESGE and VCUAM classifications. In line with our results, Heinonen et al. (2016) found that the cervical and vaginal abnormalities associated with uterine abnormalities were not adequately represented using the AFS system [17].

The VCUAM classification was rated best for adnexal and associated malformations (▶ **Fig. 2**). One reason might be that the adnexa are not listed separately in the ESHRE/ESGE classification and "associated anomalies of non-Müllerian origin," including malformations of the Wolffian ducts or of the urogenital sinus, can only be presented descriptively, as in the AFS classification. Nevertheless, Di Spiezio Sardo et al. (2015) found that the ESHRE/ESGE classification provides a description and categorization of nearly all currently known genital malformations [18].

Ludwin et al. (2015) postulated that the ESHRE/ESGE classification is more useful for classifying complex genital malformations owing to the separate assessment of the uterus, cervix, and vagina compared to the American Society for Reproductive Medicine ASRM classification. However, they also found that the ESHRE/ESGE had a significantly higher detection rate for uterine septum compared with the ASRM classification (16.9 vs. 6.9 %), which may be associated with a serious risk of over-diagnosis [19]. This could lead to increased unnecessary surgical interventions and as a result, massively higher costs for the health care system. Ouyang et al. (2018) arrived at a similar conclusion and therefore recommended using the ESHRE/ESGE classification with caution, especially with regard to hysteroscopic surgery in patients with infertility [20, 21].

This study has several limitations. First, due to the low prevalence of female genital malformations, almost half of the patients were assessed in the initiating study center 1 by a few physicians specializing in diagnosis and treatment. Half of the centers assessed 54 out of 65 patients. Fifty patients were evaluated by 4 investigators while 5 investigators evaluated only one patient. Second, the VCUAM classification was developed by one of the authors of this paper. Therefore, centers that were unbiased or co-developed the ESHRE/ESGE classification were also included in this study. Third, one aim of the EVA Study was to investigate the subjective reproducibility of three common classifications for female genital malformations in a prospective, international, multicenter study which included an opinion survey. However, further studies with higher numbers of patients with female genital malformations are needed to test the intra- and interobserver variability of these classification systems, preferably using appropriate instruments, e.g., sets of imaging studies. However, it should be mentioned that a new version of the ASRM classification was published in September 2021, indicating that the AFS classification was no longer state-of-the-art at the time of recruitment. Therefore, the EVA study was terminated early in February 2022, and further studies are needed to compare the new ASRM classification with the ESHRE/ESGE and VCUAM classifications and the existing data from the recent EVA study.

Conclusion

The prospective multicenter EVA (ESHRE/ESGE | VCUAM | AFS) study revealed that the organ-based ESHRE/ESGE and VCUAM classifications for female genital malformations appear to perform better in terms of subjective reproducibility as well as in the assessment of individual compartments than the non-organ-based AFS classification. In order to test the intra- and interobserver variability of these classification systems for female genital malformations, further prospective multicenter studies with higher numbers of patients with female genital malformations are needed, preferably using appropriate instruments, e.g., sets of imaging studies.

Supplemental Material

- Supplemental data: EVA study questionnaire.
- Supplemental **Fig. S1**: Relative frequencies of rankings of the classification methods.
- Supplemental **Fig. S2**: Relative frequencies of rankings by study center per classification method.

Funding Statement

The project was not funded, and therefore no funding sources are listed.

Attestation Statement

The subjects in this trial were not concomitantly involved in other randomized trials. Data regarding any of the subjects in the study have not been previously published unless specified. The data will be made available to the journal editors for review or query upon request.

Clinical Trial

Registration number (trial ID): DRKS00025001 | Type of Study: https://www.drks.de/ui_data_web/DrksUI.html; date of registration: 28.04.2021; date of enrollment of the first subject: 10.12.2019

Acknowledgement

The present work was performed in fulfillment of the requirements for obtaining the degree of "Dr. med." at the Friedrich-Alexander University of Erlangen-Nürnberg (FAU).

Conflict of Interest

SK, SO, RP, KR, KK, HW, PH, PT: Disclosure: none. PO: Advisory Board – PlantTec Medical GmbH. PGO: Leadership AG Kinder- und Jugendgynäkologie Germany – unpaid. RS: Editor for the Journal of Pediatric Urology – unpaid.

References

[1] Saravelos SH, Cocksedge KA, Li TC. Prevalence and diagnosis of congenital uterine anomalies in women with reproductive failure: a critical appraisal. Hum Reprod Update 2008; 14: 415–429. doi:10.1093/humupd/dmn018

[2] Byrne J, Nussbaum-Blask A, Taylor WS et al. Prevalence of Müllerian duct anomalies detected at ultrasound. Am J Med Genet 2000; 94: 9–12. doi:10.1002/1096-8628(20000904)94:19::aid-ajmg33.0.co;2-h

[3] Simón C, Martinez L, Pardo F et al. Müllerian defects in women with normal reproductive outcome. Fertil Steril 1991; 56: 1192–1193

[4] Oppelt P, Binder H, Birraux J et al. Diagnosis and Therapy of Female Genital Malformations (Part 1). Guideline of the DGGG, OEGGG and SGGG (S2k Level, AWMF Registry Number 015/052, May 2019). Geburtshilfe Frauenheilkd 2021; 81: 1307–1328. doi:10.1055/a-1471-4988

[5] Oppelt P, Binder H, Birraux J et al. Diagnosis and Therapy of Female Genital Malformations (Part 2). Guideline of the DGGG, OEGGG and SGGG (S2k Level, AWMF Registry Number 015/052, May 2019). Geburtshilfe Frauenheilkd 2021; 81: 1329–1347. doi:10.1055/a-1471-4988

[6] Toth B, Würfel W, Bohlmann M et al. Recurrent Miscarriage: Diagnostic and Therapeutic Procedures. Guideline of the DGGG, OEGGG and SGGG (S2k-Level, AWMF Registry Number 015/050). Geburtshilfe Frauenheilkd 2018; 78: 364–381. doi:10.1055/a-0586-4568

[7] Jacquinet A, Millar D, Lehman A. Etiologies of uterine malformations. Am J Med Genet A 2016; 170: 2141–2172. doi:10.1002/ajmg.a.37775

[8] Deutsche Gesellschaft für Gynäkologie und Geburtshilfe (DGGG) e.V., Österreichische Gesellschaft für Gynäkologie und Geburtshilfe (OEGGG), Schweizerische Gesellschaft für Gynäkologie und Geburtshilfe (SGGG), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF). Female genital malformation. Guideline of the DGGG, SGGG and OEGGG (S2k-Level, AWMF Registry No. 015/052, March 2020). 2020. Accessed January 12, 2023 at: http://www.awmf.org/leitlinien/detail/ll/015–052.html

[9] Acién P, Acién M, Sánchez-Ferrer M. Complex malformations of the female genital tract. New types and revision of classification. Hum Reprod 2004; 19: 2377–2384. doi:10.1093/humrep/deh423

[10] Buttram VC, jr., Gibbons WE. Müllerian anomalies: a proposed classification. (An analysis of 144 cases). Fertil Steril 1979; 32: 40–46. doi:10.1016/s0015-0282(16)44114-2

[11] Anonymous. The American Fertility Society classifications of adnexal adhesions, distal tubal occlusion, tubal occlusion secondary to tubal ligation, tubal pregnancies, müllerian anomalies and intrauterine adhesions. Fertil Steril 1988; 49: 944–955. doi:10.1016/s0015-0282(16)59942-7

[12] Acién P. Embryological observations on the female genital tract. Hum Reprod 1992; 7: 437–445. doi:10.1093/oxfordjournals.humrep.a137666

[13] Acién P, Acién MI. The history of female genital tract malformation classifications and proposal of an updated system. Hum Reprod Update 2011; 17: 693–705. doi:10.1093/humupd/dmr021

[14] Acién M, Acién P. Classification of Müllerian anomalies: Is a consensus possible? Case Rep Womens Health 2022; 34: e00413. doi:10.1016/j.crwh.2022.e00413

[15] Oppelt P, Renner SP, Brucker S et al. The VCUAM (Vagina Cervix Uterus Adnex-associated Malformation) classification: a new classification for genital malformations. Fertil Steril 2005; 84: 1493–1497. doi:10.1016/j.fertnstert.2005.05.036

[16] Grimbizis GF, Gordts S, Di Spiezio Sardo A et al. The ESHRE/ESGE consensus on the classification of female genital tract congenital anomalies. Hum Reprod 2013; 28: 2032–2044

[17] Heinonen PK. Distribution of female genital tract anomalies in two classifications. Eur J Obstet Gynecol Reprod Biol 2016; 206: 141–146. doi:10.1016/j.ejogrb.2016.09.009

[18] Di Spiezio Sardo A, Campo R, Gordts S et al. 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 2015; 30: 1046–1058

[19] Ludwin A, Ludwin I. Comparison of the ESHRE-ESGE and ASRM classifications of Müllerian duct anomalies in everyday practice. Hum Reprod 2015; 30: 569–580. doi:10.1093/humrep/deu344

[20] Ludwin A, Ludwin I, Coelho Neto MA et al. Septate uterus according to ESHRE/ESGE, ASRM and CUME definitions: association with infertility and miscarriage, cost and warnings for women and healthcare systems. Ultrasound Obstet Gynecol 2019; 54: 800–814. doi:10.1002/uog.20291

[21] Ouyang Y, Yi Y, Gong F et al. ESHRE-ESGE versus ASRM classification in the diagnosis of septate uterus: a retrospective study. Arch Gynecol Obstet 2018; 298: 845–850. doi:10.1007/s00404-018-4878-2