Human Reproduction Open, Vol.00, No.0, pp. 1–10, 2020
doi:10.1093/hropen/hoaa053

human reproduction open

ORIGINAL ARTICLE

EQUSUM: Endometriosis QUality and grading instrument for SUrgical performance: proof of concept study for automatic digital registration and classification scoring for r-ASRM, EFI and Enzian

J. Metzemaekers [1], P. Haazebroek[2], M.J.G.H. Smeets[3], J. English[3], M.D. Blikkendaal[1], A.R.H. Twijnstra[1], G.D. Adamson[4], J. Keckstein[5,6], and F.W. Jansen [1,7,*]

[1]Department of Gynaecology, Leiden University Medical Center, Leiden, the Netherlands [2]Institute of Psychology, Leiden University, Leiden, the Netherlands [3]Department of Gynaecology, Haaglanden Medisch Centrum-Bronovo, Den Haag, the Netherlands [4]Equal3 Fertility, Cupertino, CA, USA [5]Stiftung Endometrioseforschung (SEF), Westerstede,Germany [6]Gynecological Clinic Drs. Keckstein, Villach, Austria [7]Department of Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands

*Correspondence address. Department of Gynaecology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, the Netherlands. E-mail: f.w.jansen@lumc.nl http://orcid.org/0000-0002-5300-1694

Submitted on May 17, 2020; resubmitted on September 22, 2020; editorial decision on October 8, 2020

STUDY QUESTION: Is electronic digital classification/staging of endometriosis by the EQUSUM application more accurate in calculating the scores/stages and is it easier to use compared to non-digital classification?

SUMMARY ANSWER: We developed the first digital visual classification system in endometriosis (EQUSUM). This merges the three currently most frequently used separate endometriosis classification/scoring systems (i.e. revised American Society for Reproductive Medicine (rASRM), Enzian and Endometriosis Fertility Index (EFI)) to allow uniform and adequate classification and registration, which is easy to use. The EQUSUM showed significant improvement in correctly classifying/scoring endometriosis and is more user-friendly compared to non-digital classification.

WHAT IS KNOWN ALREADY: Endometriosis classification is complex and until better classification systems are developed and validated, ideally all women with endometriosis undergoing surgery should have a correct rASRM score and stage, while women with deep endometriosis (DE) should have an Enzian classification and if there is a fertility wish, the EFI score should be calculated.

STUDY DESIGN, SIZE, DURATION: A prospective endometriosis classification proof of concept study under experts in deep endometriosis was conducted. A comparison was made between currently used non-digital classification formats for endometriosis versus a newly developed digital classification application (EQUSUM).

PARTICIPANTS/MATERIALS, SETTING, METHODS: A hypothetical operative endometriosis case was created and summarized in both non-digital and digital form. During European endometriosis expert meetings, 45 DE experts were randomly assigned to the classic group versus the digital group to provide a proper classification of this DE case. Each expert was asked to provide the rASRM score and stage, Enzian and EFI score. Twenty classic forms and 20 digital forms were analysed. Questions about the user-friendliness (system usability scale (SUS) and subjective mental effort questionnaire (SMEQ)) of both systems were collected.

MAIN RESULTS AND THE ROLE OF CHANCE: The rASRM stage was scored completely correctly by 10% of the experts in the classic group compared to 75% in the EQUSUM group ($P < 0.01$). The rASRM numerical score was calculated correctly by none of the experts in the classic group compared with 70% in the EQUSUM group ($P < 0.01$). The Enzian score was correct in 60% of the classic group compared to 90% in the EQUSUM group ($P = 0.03$). EFI scores were calculated correctly in 25% of the classic group versus 85% in

the EQUSUM group ($P < 0.01$). Finally, the usability measured with the SUS was significantly better in the EQUSUM group compared to the classic group: 80.8 ± 11.4 and 61.3 ± 20.5 ($P < 0.01$). Also the mental effort measured with the SMEQ was significant lower in the EQUSUM group compared to the classic group: 52.1 ± 18.7 and 71.0 ± 29.1 ($P = 0.04$). Future research should further develop and confirm these initial findings by conducting similar studies with larger study groups, to limit the possible role of chance.

LIMITATIONS, REASONS FOR CAUTION: These first results are promising, however it is important to note that this is a preliminary result of experts in DE and needs further testing in daily practice with different types (complex and easy) of endometriosis cases and less experienced gynaecologists in endometriosis surgery.

WIDER IMPLICATIONS OF THE FINDINGS: This is the first time that the rASRM, Enzian and EFI are combined in one web-based application to simplify correct and automatic endometriosis classification/scoring and surgical registration through infographics. Collection of standardized data with the EQUSUM could improve endometriosis reporting and increase the uniformity of scientific output. However, this requires a broad implementation.

STUDY FUNDING/COMPETING INTEREST(S): To launch the EQUSUM application, a one-time financial support was provided by Medtronic to cover the implementation cost. No competing interests were declared.

TRIAL REGISTRATION NUMBER: N/A.

Key words: endometriosis/ operative reporting/ surgery/ rASRM/ EFI/ Enzian/ classification/ electronic

WHAT DOES THIS MEAN FOR PATIENTS?

This proof of principle study looks at whether digital endometriosis classification with the EQUSUM application is more precise compared to the current non-digital classification.

Endometriosis classification remains a difficult topic, while there is no system that covers the disease completely. The revised American Society for Reproductive Medicine classification system is mainly useful for less aggressive endometriosis, while the Enzian classification system is more suitable for deep endometriosis and the Endometriosis Fertility Index score is not really a endometriosis classification system, but predicts fertility chances after surgery.

It is challenging for the clinician to know all these classification systems (with exception rules), which makes it difficult to incorporate it in the daily clinical setting. Therefore, it is desirable to make endometriosis classification more accessible and easier to use.

The researchers concluded with their study that digitalization of endometriosis classification has the potential to make it more accurate and uniform and easier to use.

Introduction

Both patients and healthcare providers criticize current endometriosis classification systems, because there is poor correlation of the extent of the disease with symptoms and prognosis (Rolla, 2019). This raises the question: why is proper classification important? And how come that the current classification translates poorly to clinical outcomes?

Classification is important, because the goal of classification is to provide a common language for healthcare providers to communicate, monitor and report diseases (Jutel, 2011). This way we are able to standardize and communicate about the same disease between healthcare workers, hospitals and countries. Furthermore, it enables correct collection and storage of data for scientific research.

For endometriosis, the World Endometriosis Society (WES) stated:

> 'Until better classification systems are validated, all women with endometriosis undergoing surgery should have an rASRM score and stage completed, women with deep endometriosis should have an Enzian classification completed, and women for whom fertility is a future concern should have an EFI score completed, and documented in the medical/surgical notes' (strong GPP = good practice point) (Johnson *et al.*, 2017).

The WES and World Endometriosis Research Foundation (WERF) are aware of the fact that three separate systems are not ideal, but until and if a comprehensive classification system that covers endometriosis in general, deep endometriosis (DE) and fertility outcomes are created, this is the best that can be done at this time. However, digitalization of the systems in use, could be a promising solution to overcome these current issues stated above.

Usability of the current systems: rASRM, Enzian and EFI

The revised American Society for Reproductive Medicine (rASRM) classification system (1997) is the most widely adopted and claims to be relatively easy to use (Haas *et al.*, 2013a). However, there are exception rules in the original description which are only known after reading the article very meticulously. Furthermore, with the rASRM, it is only possible to describe the peritoneum, ovaries, tubes, pouch of Douglas and uterus. It is well known that endometriosis does not stay within the anatomical boundaries of the internal reproductive organs. Therefore, classification limited to these structures will not cover the disease *in toto*. The rASRM total points lead to stage I–IV (I = 1–5,

II = 6–15, III = 16–40, IV = >40), but no information on the location of the lesion is provided. Furthermore, the rASRM is often used to classify DE (Timoh *et al.*, 2018; Leonardi *et al.*, 2019; Abrão *et al.*, 2020), while it inherently cannot account for DE. The rASRM is not designed to classify endometriosis lesions in organs such as the rectum, bladder, ureter and other DE structures and is therefore not the classification system of choice in DE.

The Enzian (Keckstein *et al.*, 2003) classification is specifically developed for DE. It is mostly used in German-speaking countries, has relatively poor international acceptance and is seen as a more complex/difficult system (Haas *et al.*, 2013a). The advantages are well recognized, because it scores not only different regions of the pelvis but also provides information about anatomical location. Such details are useful in both research and clinical care. However, the Enzian provides, up till now, no information on the tubes, ovaries and peritoneum. The *Stiftung Endometriose Forschung* (SEF) is working on an update and will incorporate these structures in a revised Enzian scoring system #Enzian.

The Endometriosis Fertility Index (EFI) scoring system (Adamson and Pasta, 2010) provides fertility chances after surgery and uses two rASRM scores, total lesion score and total score (lesions plus adhesions). The EFI has been validated in over 24 studies (Cook and Adamson, 2013; Tomassetti *et al.*, 2013; Kim *et al.*, 2019) and provides correct natural fertility chances after surgery. The EFI score defines four treatment levels (I–IV), with recommendations ranging from 'attempt non-ART conception for at least 1 year' to 'refer to ART centre for IVF' (Cook and Adamson, 2013).

Is it time for a new classification system?

Since the aetiology and pathophysiology of endometriosis is not completely revealed, the diagnosis and stage is not uniform and subsequent treatment regimens vary. Unfortunately to date, there is no system that includes all: staging of general endometriosis, DE and fertility prognosis. Therefore, it is arguable whether we should introduce a new system if we are not able to uniformly use the existing classification systems. New classification systems have frequently been and are being developed (Koninckx *et al.*, 2011; Khazali, 2016), but a single useful and comprehensive system has not been created. Endometriosis remains a complex disease with divergent phenotypes and features and is therefore challenging to capture in one classification system.

Furthermore, it has to be addressed that better description of endometriosis lesions may not automatically means improvement of care. The natural history and course of endometriosis is difficult to register and is subject to multifactorial influences (hormone levels, medication use, age) which has effect on the progression of the disease. Compared to oncologic systems, endometriosis classification is more complex due to the lack of a single hard outcome measure (e.g. survival in oncological versus endometriosis recurrence with several definitions and monitoring difficulties). It should be realized that we can't classify in the same way as oncological diseases, and more effort is needed to create uniform outcome measures which enable research and, in the end, better treatment for endometriosis patients.

Uniformity in surgical documentation and registration in DE

The publication Consensus on Recording Deep Endometriosis Surgery (CORDES) (Vanhie *et al.*, 2016) formulated Deep Endometriosis Surgical Sheets (DESS) for uniform registration in DE surgery. These sheets are composed from consensus of international experts, based on a systematic review of the literature. The use of these surgical sheets combined with correct staging could be of importance in scientific research, especially since more countries such as the UK (Saridogan and Byrne, 2013) and German-speaking countries (Ebert *et al.*, 2013) are formulating and implementing guidelines for centralization of endometriosis care. This will only succeed if there is standardization and harmonization of research and terminology.

Digitalization and technical development of the EQUSUM

Digitalization in healthcare has the potential to improve it (Gellerstedt, 2016). Therefore, digitalizing current classification systems could help physicians classify endometriosis with require less intellectual effort, hold extensive knowledge of systems and exception rules, create fewer errors and take less time. It facilitates uniform and automatic database formation. Based on the CORDES statements, we incorporated a significant part of the DESS into the EQUSUM application. Based on the recommendation of the WES (Johnson *et al.*, 2017), we processed the rASRM, Enzian and EFI into a web-based system, without changing these current classification systems. In this application, it is also possible to register the postoperative complications as defined by Clavien–Dindo (Dindo *et al.*, 2004). This enables the possibility to link complications to different surgical techniques, which could provide more insight in DE surgical treatments. For example, all surgical techniques (type of anastomosis, sewing technique etc.) are uniform documented, and with robust data it would be possible to detects trends in certain techniques with certain complications (e.g. (fictionally) more anastomosis leakage with hand-sewn end-to-end anastomosis compared to side-to-end anastomosis).

Our web-based system is called EQUSUM, Endometriosis QUality and grading instrument for SUrgical perforMance. The advantage of the application is that it generates automatic calculated classification and registration in general endometriosis and DE on anatomical pictures. The calculated EFI is incorporated for fertility chances after surgery. Neurological involvement and nerve sparing surgical techniques can be documented in one database with uniform surgical data. Finally, the application incorporates automatically the usual and complex classification tables, figures and calculations, including the exception rules. The EQUSUM is a digital entry system that automates calculation of three previously established and recommended scoring systems. Users can register for free on www.equsum.org. Surgical procedures can be filled in, and free text or other traceable patient characterizes can be documented, making it safe to use (based on data experts review). The endometriosis lesions can be assigned on anatomical pictures, which automatically calculates an rASRM score and stage, Enzian and EFI score (if applicable). This classification can be documented in the patient's file. Surgeons who register their procedures can freely get their registered surgical data (surgical procedure, techniques used, classification, surgical outcomes etc.) at the end of each year (syntax for SPSS will be given as well, making data analysis easy).

Our research aim is to find a solution for the challenging problem to obtain a correct endometriosis classification and staging. Therefore we compare the separate (on paper) classic systems rASRM, Enzian and EFI to the digital classification application EQUSUM. The endpoints of this study will be the correct staging/scoring and user-friendliness.

Materials and methods

We performed a prospective proof of concept study with expert gynaecologist in endometriosis, from different international centres. Data were collected from 1 January through 30 April 2019. The currently used non-digital ('classic') classification system are compared to the outcome of the digital EQUSUM system. The latter, web-based system was developed by our institute in close collaboration with the Department of Psychology and Information Technology of the Leiden University and meets all required standards for data security. This collaboration gained broad experience with previous scientific projects (Driessen *et al.*, 2016). Recruitment of participants took place at European congresses for endometriosis. Only participants with known experience in endometriosis were approached for this inquiry. We also collected baseline information about their experience. The inquiry was randomly assigned, subsequently handing out (1:1) a digital form followed by a non-digital test and so on. On the digital form, instructions were given to create an EQUSUM account (www.equsum.org) and basic instructions were provided on how to fill in the information from the operative case. Informed consent was obtained from each participating physician. Self-reported general data about experience in endometriosis surgery, DE and surgical cases on average per year were collected.

The inclusion criteria were experts in endometriosis surgery. We considered >3 years of specific surgical experience in endometriosis as experienced.

Exclusion criteria were forms that did not follow the given instructions correctly e.g. filling in a non-digital form while a digital form was handed out and incomplete forms regarding the classification.

Case

A hypothetical patient case was created (Supplementary Fig. S1): 34-year-old woman, para 1, and trying to get pregnant for one year without success. She has a diagnosis of DE of the bladder with complaints of pain and scheduled for surgery. All exception rules of the classification systems were included in this case.

Outcomes and definitions

The primary outcome was the correct staging/classification of the hypothetical patient.

Secondary outcomes were the application of the exceptions rules of each classification system and user-friendliness measured by the Subjective Mental Effort Questionnaire (SMEQ) (Bevan, 1995) and system usability scale (Brooke, 1996). In detail, exception rules for the rASRM consisted of assignment of 16 points to the fallopian tube with a complete enclosure of the fimbria by adhesions and multiplied points on the remaining tube and ovary due to the missing adnexa (1997). For the Enzian classification bilateral involvement needs to be assigned with two characters (Haas *et al.*, 2013b). Furthermore, negative findings in compartment A, B and C need to be described as well (Haas *et al.*, 2013a,b). With the EFI, the Least Function (LF) needs to be doubled if an ovary is missing (Adamson and Pasta, 2010).

The SMEQ has a scale from 0 (not at all hard to do) to 150, whereby scores above 110 means that it is tremendously hard to do. The system usability scale score varies from 0 (not at all user-friendly) to 100 (user-friendly), and can be divided in the following categories; <51 awful, 51–68 poor, 68 okay, 68–80.3 good, >80.3 excellent.

Correct classification of the hypothetical patient case

rASRM

rASRM endometriosis score is 6 points of the peritoneum and 4 points of the ovary (×2 because the missing ovary), resulting in $6+8$ $(4 \times 2) = 14$ points.

The adhesions score is 8 points for the left ovary (×2 because of the missing ovary) and 8 points for the right tube (assigned to 16 due to complete enclosure of the fimbriae), resulting in a total of 32 adhesion points.

Together these result in a total rASRM score of $14+32=46$ points, stage IV.

Enzian

The correct notation for the Enzian is A0BB2C0, FB. Because of bilateral involvement of compartment B we only take the highest score. Shorter descriptions (used in clinical setting), were also stated correctly (B1B2, FB or B2, FB).

EFI

The LF for the EFI results in taking the lowest score of both sites. The right side is 0 due to the missing ovary. On the left side, the lowest score is the left fallopian tube scored moderate 2 points, resulting in a total LF of 4 points, because the site with the ovary must be doubled (2×2).

The correct historical factor is 5. Age being under 35 gives 2 points. Years infertile is <3 years also 2 points. And prior pregnancy gives one point. These results in 5 points in total.

The surgical factor is 2 points resulting from the LF (4–6 is moderate, 2 points). rASRM score endometriosis is lower than 16, giving 1 point. And the total rASRM score <71 gives 1 point. Surgical factor in total is therefore 4 points.

EFI score is historical (5 points) + surgical (4 points) = EFI score of 9.

Statistical analysis

IBM SPSS version 25.0 for Windows was used for our analysis and we used the Shapiro–Wilk test to evaluate the distribution of the data. Data are presented as mean ± SD or median (with interquartile range) for normally distributed or skewed data, respectively. We used Student's *t*-test for normally distributed data and Mann–Whitney *U* test for skewed data. Categorical data were compared using χ^2 test. Odds ratios (ORs) were calculated from univariate logistic regression model. Given the overarching aim of this proof of concept study (explore the correctness and usability of digital classification), a power calculation was not conducted, which is common practice amongst pilot studies (Billingham *et al.*, 2013). We choose for 20 inclusions per

arm. We considered a two-tailed *P*-value of <0.05 as statistically significant.

Ethical approval

Ethical approval was given by the Medical Ethics Committee of the Leiden University Medical Center (LUMC) (G20.019). Informed consent was obtained by each participating gynaecologist.

Results

Baseline characteristics of participating experts in deep endometriosis

In total, 45 participants were recruited, and five were excluded with the following reasons: two because of not being expert in endometriosis surgery, two because of missing data about experience and one due to not following the instructions (filled in classification on paper while it was a digital form). This resulted in 20 experts completing the non-digital classification format test (classic version) and 20 completing the digital (EQUSUM) version.

Baseline characteristics are presented in Table I. There were no statistical differences in the median (interquartile range) experience (years) in general endometriosis between the classic and the EQUSUM groups, respectively 18 (7–30) and 18 (6–23) ($P = 0.32$). Also, no statistical difference was found in the amount of surgical DE cases per year, respectively a median of 60 (25–113) and 30 (20–49) ($P = 0.08$). A statistical difference was found in general surgical endometriosis cases per year, a median of 150 (65–313) in the classic group and 55 (25–100) in the EQUSUM group ($P < 0.01$).

Correct use of the rASRM, Enzian and EFI scores

Statistical differences were found on all primary outcome measures between the classic group and EQUSUM (Table II). In the classic group, 10% scored the rASRM stage correctly, compared to 75% in the EQUSUM group ($P < 0.01$, OR: 27.0; 95% CI: 4.6–159.7). Considering the rASRM score, none of the experts in the classic group but 70% in the EQUSUM group scored this completely correctly ($P < 0.01$).

The Enzian classification showed a statistical differences: the classic group was correct with 60% of expert, compared to 90% in the EQUSUM group ($P = 0.03$, OR: 6.0; 95% CI: 1.1–33.3). For the EFI score, 25% were correct in the classic group, whereas in the EQUSUM group, 85% scored this correctly ($P < 0.01$, OR: 17.0; 95% CI: 3.5–83.4).

Exceptions rules applied

For the rASRM classification, the exception rule when the fimbriated end of the fallopian tube is completely enclosed (i.e. change the point assignment to 16) was used in 90% in the classic group compared to 85% in the EQUSUM group (Table III). The exception rule of only one adnexa was applied in one case (5%) in the classic group, whereas it was automatically calculated in 95% in the EQUSUM group. No discrepancy was found in the EQUSUM group between the rASRM score and stage, this discrepancy occurred in the classic group in 5% of the cases. (This means assigning the wrong stage in relation to the rASRM score.)

In the classic group, seven versions of the Enzian were given and three were given in the EQUSUM group. In the classic group, the majority used the B2, FB nomenclature (50%). The terminology B1B2 for compartment B was also considered correct; this was used in two forms (10%) resulting in 60% correct Enzian classification. In the EQUSUM group 85% generated a correct Enzian classification of A0, BB2, C0, FB.

Considering the EFI components in the classic group, the LF was reported correctly in 30%, historical factor correct in 95%, LF was doubled in 45% of the cases, surgical factors were filled in correctly in 25% and the endometriosis rASRM score was correctly used in 40%. In the EQUSUM group, the LF was correct in 95%, historical factor was correct in 95%, LF was doubled in 95%, surgical factors were correct in 95% and there was correct use of the endometriosis rASRM score in 95%.

User convenience

The usability scored with the system usability scale was significantly different between the classic group (n = 17) and EQUSUM group (n = 16), with respectively a mean of 61.3 ± 20.5 and 80.8 ± 11.4 ($P < 0.01$). In Table IV, the system usability scale score is divided in categories from excellent to awful. In the classic group, 15% thought the classification was excellent, whereas in the EQUSUM group this was 45%. None of the participants scored the EQUSUM as awful, compared to 30% in the classic group.

The mental effort measured with the SMEQ showed that the EQUSUM scored lower (less effort) compared to the classic group. In the classic group (n = 19), a mean of 71.0 ± 29.1 was found and in the EQUSUM group (n = 14), a mean of 52.1 ± 18.7 ($P = 0.04$).

Table I Baseline characteristics of experts in endometriosis.

Characteristics (self-reported)	Classic (n = 20)	EQUSUM (n = 20)	*P*-value
Experience general endometriosis, median (Q1–Q3) years	18 (7–30)	18 (6–23)	0.32
Surgical cases general endometriosis, median (Q1–Q3) case/year*	150 (65–313)	55 (25–100)	<0.01
Surgical cases deep endometriosis, median (Q1–Q3) case/year**	60 (25–113)	30 (20–49)	0.08

*Surgical cases general endometriosis; missing values: 2 (classic).
**Surgical cases deep endometriosis; missing values: 3 (classic).
Interquartile range is reported with Q1 and Q3.

Table II Primary outcome measures.

	Classic (n = 20)	EQUSUM (n = 20)	Univariable*Odds ratio (95% CI)	*P*-value
rASRM score right, n (%)	0 (0)	14 (70)	–	<0.01
rASRM stage right, n (%)	2 (10)	15 (75)	27.0 (4.6–159.7)	<0.01
Enzian score right, n (%)	12 (60)	18 (90)	6.0 (1.1–33.3)	0.03
EFI score right, n (%)	5 (25)	17 (85)	17.0 (3.5–83.4)	<0.01

*Odds ratio not possible to calculate for rASRM score because of zero in the equation.
EFI, Endometriosis Fertility Index; rASRM, revised American Society for Reproductive Medicine.

Table III Secondary outcome measures, exception rules of rASRM, Enzian and EFI as scored correctly by number of participants.

	Classic (n = 20)	EQUSUM (n = 20)
rASRM classification, n (%)		
Exception rule fimbria used (assign to 16)	18 (90)	17 (85)
Exception rule ovary used (double points)	1 (5)	19 (95)
Score/stage error*	1 (5)	0 (0)
Enzian versions classic, n (%)		
B2, FB**	10 (50)	–
AB2FB	1 (5)	–
B1(li), B2(re), FB	3 (15)	–
B1B2, FB**	2 (10)	–
B2rB1l, FB	2 (10)	–
AB2dextC0 F(B)	1 (5)	–
2, B2, FB	1 (5)	–
Enzian versions EQUSUM, n (%)		
A2, BB2, C0, FB	–	2 (10)
A0, B0, C0	–	1 (5)
A0, BB2, C0, FB**	–	17 (85)
EFI score, n (%)		
LF right calculation	6 (30)	19 (95)
Historical factor right calculation	19 (95)	19 (95)
LF doubled	9 (45)	19 (95)
Surgical factor right use	5 (25)	19 (95)
Right use of endometriosis score rASRM	8 (40)	19 (95)

*Assigning the wrong stage in relation to the rASRM score.
**Correct Enzian score.
LF, Least Function; rASRM, revised American Society for Reproductive Medicine.

Table IV User convenience.

Characteristics	Classic	EQUSUM	*P*-value
System Usability Scale (SUS), mean (SD)*	61.3 (20.5)	80.8 (11.4)	<0.01
Excellent (%)	15%	45%	
Good (%)	20%	20%	
Poor (%)	20%	15%	
Okay (%)	0%	0%	
Awful (%)	30%	0%	
Not respond (%)	15%	20%	
Subjective Mental Effort Questionnaire (SMEQ), mean (SD)**	71.0 (29.1)	52.1 (18.7)	0.04

*SUS missing values: 3 (classic), 4 (EQUSUM)—<51 awful, 51–68 poor, 68 okay, 68–80.3 good, >80.3 excellent.
**SMEQ; missing values: 1 (classic), 6 (EQUSUM)—scale from 0 to 150, 0 means not at all hard to do, above 110 means tremendously hard to do.

Discussion

The results of this proof of concept study with the EQUSUM application shows that classification of DE by endometriosis experts was significantly more correct when using this digital system compared to the currently in use, classic classification on paper. With the EQUSUM, clinicians do not have to consider the complex exception rules, because it is automatically generated and processed in the application through infographics. This has the potential to improve classification of (deep) endometriosis.

With all the current exception rules (Table V) in endometriosis classification, clinicians have to study the original papers carefully to find out about these rules. It is well known that in daily practice, most of the clinicians don't use the rASRM table for classification, but stage by their clinical view/intuition.

It is unrealistic to expect that all gynaecologists can apply these rules to three different classification systems flawlessly. Furthermore, we have to consider the user-friendliness of classification systems. If a system is too labour-intensive, it is likely and understandable that doctors will not use the system. In our study, we found that the majority scored that the classic classification as awful, while in the EQUSUM group the majority scored it as excellent and none scored it awful. The digital classification system seems very easy to use, due to visual advantages with anatomical pictures for the classification (Fig. 1, print screens of EQUSUM interface; an extensive overview is found on www.equsum.org). Using so-called infographics instead of words/tables, provides easy and quick understanding (McCrorie *et al.*, 2016). Another advantage of the digital version is that if new rules or updates

Table V Exception rules current classification/scoring systems.

rASRM	1. If the fimbriated end of the fallopian tube is completely enclosed, change the point assignment to 16 (1997). 2. In those patients with only one adenexa, points applied to disease of the remaining tube and ovary should be multiplied by two (1997). 3. The severity of the endometriosis or adhesions should be assigned the highest score only for peritoneum, ovary, tube or culdesac (1997).
Enzian	1. Bilateral involvement noted with two letter for B (BB) and FU (FUU) (Haas *et al.*, 2013b) 2. Negative findings for compartment ABC are noted, A0B0C0 (Haas *et al.*, 2013a,b)
EFI	1. If an ovary is absent on one side, the LF score is obtained by doubling the lowest score on the side with the ovary (Adamson and Pasta, 2010).

EFI, Endometriosis Fertility Index; rASRM, revised American Society for Reproductive Medicine.

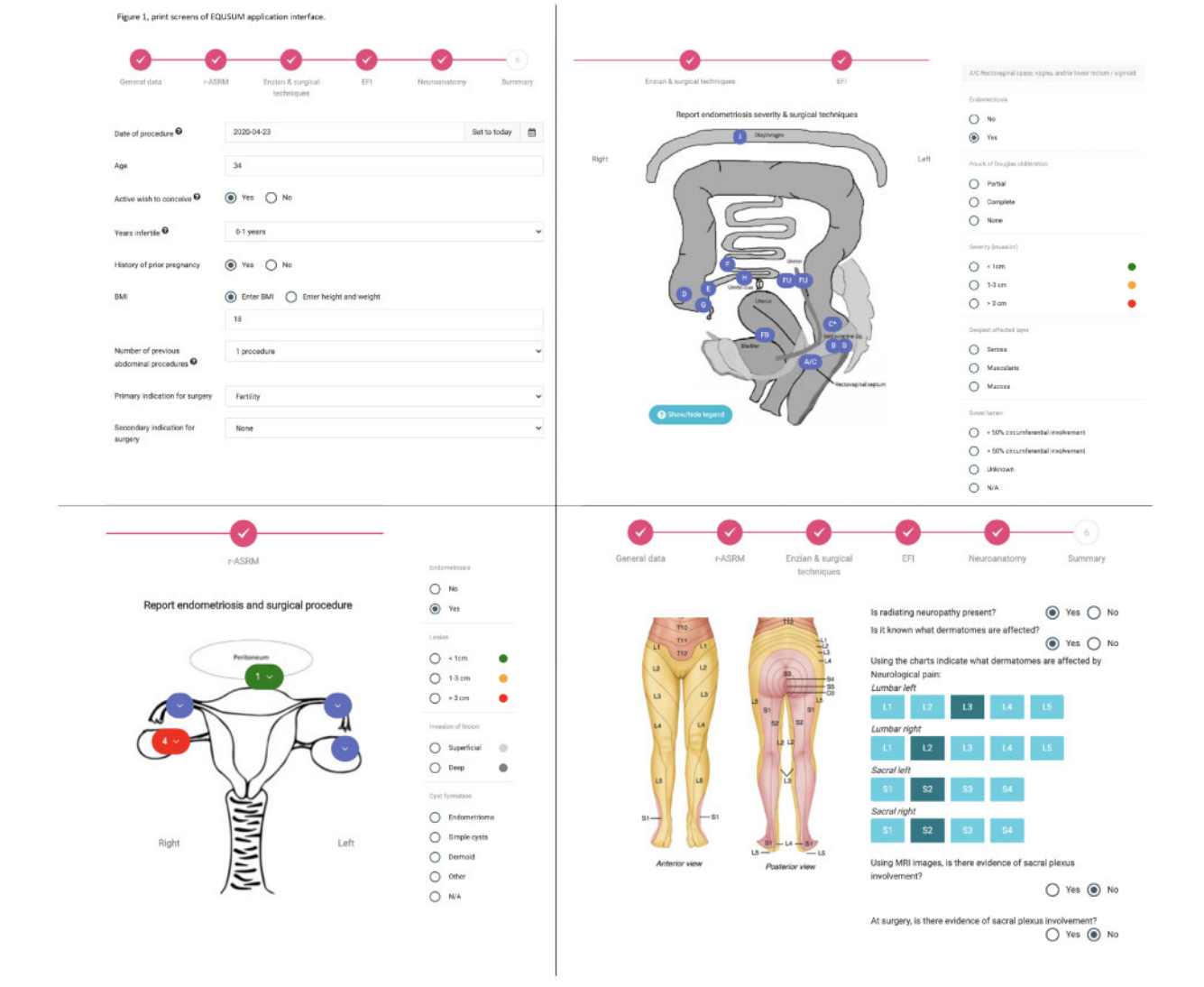

Figure 1. Print screens of EQUSUM interface.

appear for classification, it can easily be adjusted making it an optimal and dynamic digital system.

It has to be noted that the EQUSUM is comparable with a calculator, wrong input does not give correct output, although the calculation behind it, is correct. Inherently this is also a limitations of the EQUSUM application. Even if all exception rules are built-in and classification can be done on clear anatomical pictures, digital classification is still not 100% flawless. Classification in the EQUSUM application fails when, the general data is not filled in correctly; e.g. if one ovary is missing and this is not entered correctly in the general data, the exception rule for a missing ovary will not be applied and therefore not be calculated in the rASRM and EFI score.

Furthermore, we tested this proof of concept study on experts in deep endometriosis and gave a difficult DE case to analyse. Further research with different physicians and diverse patient cases with less exception rules can make the system even more robust.

Another main issue that needs to be addressed is that all classification systems for endometriosis did not prove to be predictive in terms of likelihood of conception (except EFI), pain and lesion recurrence and, most importantly, the response to surgical and conservative treatments. Therefore, the question remains whether merging three classification systems in a single application will result in improvement of the points addressed. However, creating a uniform 'language', whereby doctors can communicate and use uniform data (created by EQUSUM), enables comparison and more reliable research outcomes.

With the development of the EQUSUM, we choose for classification based on surgical findings. However, with the current trend in the use of ultrasound and MRI detection and classification of the disease, we have to consider that image guided classification can only be performed with the Enzian. No rASRM classification can be made purely on ultrasound (as adhesions are difficult or not visible) and therefore it is additionally not possible to calculate an EFI score (which is based on postoperative findings and related to rASRM score). In future versions of the EQUSUM, it would be possible to incorporated Enzian scores based on radiological findings too.

Clinical relevance

The main achievements, including contributions to the field can be recapped as follows: (i) The current classification/scoring systems on a hypothetical DE case is difficult to score correctly due to the (not applied) exception rules. (ii) Current classification systems are challenging and prone to errors and makes comparison of data on (inter)national level difficult in clinical and research settings. (iii) Digitalization has the potential to improve automatically correct scoring and is more user-friendly. (iv) Taken these points together, it is important to consider how to interpret the current classification and literature in relation to prognosis and future patient treatments.

In the literature, a poor clinical correlation between rASRM classifications has been reported (Vercellini *et al.*, 1996; Johnson *et al.*, 2017). For the Enzian classification, future research is needed to confirm the correlation with clinical findings and complaints (Montanari *et al.*, 2019) and the use of preoperatively classification (Burla *et al.*, 2019). Our data showed that a significant percentage of experts do not apply exception rules correctly, resulting in erroneous DE classification and scoring. We also showed that incorrect use of the rASRM has a direct influence on the EFI score, with a significant clinical impact on fertility chances after surgery. This undoubtedly results in varying information provision to couples trying to get pregnant.

Goal of EQUSUM

The goal of the EQUSUM application is consistent with WES statements and CORDES to create uniform surgical (and complication) registration and classification/scoring for endometriosis. This enables clear and unambiguous communication between healthcare providers and scientific researchers. We did not create a new classification system, only simplified the use of current recommended systems in a visual digital automatic form. We provided a clear overview of the pros and cons of the rASRM, Enzian, EFI and EQUSUM systems (Supplementary Table SI). Because classification is subject to constant changing rules and new versions, one digital system makes it easy to update the system with these new insights. The German endometriosis research foundation (SEF) extended the Enzian classification in early 2019 (Stiftung Endometriose Forschung, 2019). In addition, the locations on the peritoneum and ovary and the status of the tubes are considered. Digitalization has the potential that all users apply the same system with the same output and are less prone to different versions or interpretations of the same classification system. This is the first step in a process of combining the three systems in one functional and easy to use system. However, this is only possible if there is a significant amount of data to be generated by the EQUSUM whereby we can link and compare all three systems on similarities, differences, strengths and limitations.

Furthermore, we postulate that general consensus for the EQUSUM application needs to be discussed with other EQUSUM users worldwide in close collaboration with the WES, WERF, SEF, European Endometriosis League and ASRM.

Conclusion

This is the first attempt at creating a digital visual and automatic scoring system for optimal classification and registration of (deep) endometriosis. This system improves correct classification/scoring of the currently recommended rASRM, Enzian and EFI score. Our proof of concept study showed that not all exceptions rules are applied by expert endometriosis surgeons, leading to incorrect scoring. This has an impact on clinically relevant outcomes such as incorrect staging and different EFI score outcomes. The EQUSUM showed significant improvement in classifying endometriosis and is more user-friendly compared to non-digital classification. Herewith we have set the first step for a worldwide web-based dynamic registration and classification/scoring system for (deep) endometriosis.

Supplementary data

Supplementary data are available at *Human Reproduction Open* online.

Authors' roles

The research was designed and written by J.M, A.T and F.J. The development of the EQUSUM application was performed by J.M and P.H.

Inclusion of participants was done by J.M and J.K. Statistical analysis was performed by J.M. All other authors contributed equally to the writing of the manuscript (i.e. critically reviewed and improved the draft version). Final approval of the version to be published was done by the last author F.J.

Funding

To launch the EQUSUM application, a one-time financial support was provided by Medtronic.

Conflict of interest

None declared.

References

Abrão MS, Andres MP, Barbosa RN, Bassi MA, Kho RM. Optimizing perioperative outcomes with selective bowel resection following an algorithm based on preoperative imaging for bowel endometriosis. *J Minim Invasive Gynecol* 2020;**27**:883–891.

Adamson GD, Pasta DJ. Endometriosis fertility index: the new, validated endometriosis staging system. *Fertil Steril* 2010;**94**: 1609–1615.

Bevan N. Measuring usability as quality of use. *Software Qual J* 1995; **4**:115–130.

Billingham SA, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. *BMC Med Res Methodol* 2013;**13**:104.

Brooke J. SUS-A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL (eds). Usability Evaluation in Industry. London: Taylor & Francis, 1996, 189–194.

Burla L, Scheiner D, Samartzis EP, Seidel S, Eberhard M, Fink D, Boss A, Imesch P. The ENZIAN score as a preoperative MRI-based classification instrument for deep infiltrating endometriosis. *Arch Gynecol Obstet* 2019;**300**:109–116.

Cook AS, Adamson GD. The role of the endometriosis fertility index (EFI) and endometriosis scoring systems in predicting inferility outcomes. *Curr Obstet Gynecol Rep* 2013;**2**:186–194.

Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;**240**:205–213.

Driessen SR, Van Zwet EW, Haazebroek P, Sandberg EM, Blikkendaal MD, Twijnstra AR, Jansen FW. A dynamic quality assessment tool for laparoscopic hysterectomy to measure surgical outcomes. *Am J Obstet Gynecol* 2016;**215**:754.e1-754.e8.

Ebert AD, Ulrich U, Keckstein J, Muller M, Schindler AE, Sillem M, Tinneberg HR, De Wilde RL, Schweppe KW; on behalf of the Endometriosis Research Foundation, and the European Endometriosis League. Implementation of certified endometriosis centers: 5-year experience in German-speaking Europe. *Gynecol Obstet Invest* 2013;**76**:4–9.

Gellerstedt M. The digitalization of health care paves the way for improved quality of life. *J Syst Cybernet Inform* 2016;**14**:1–10.

Haas D, Chvatal R, Habelsberger A, Schimetta W, Wayand W, Shamiyeh A, Oppelt P. Preoperative planning of surgery for deeply infiltrating endometriosis using the ENZIAN classification. *Eur J Obstet Gynecol Reprod Biol* 2013b;**166**:99–103.

Haas D, Shebl O, Shamiyeh A, Oppelt P. The rASRM score and the Enzian classification for endometriosis: their strengths and weknesses. *Acta Obstet Gynecol Scand* 2013a;**92**:3–7.

Johnson NP, Hummelshoj L, Adamson GD, Keckstein J, Taylor HS, Abrao MS, Bush D, Kiesel L, Tamimi R, Sharpe-Timms KL *et al.* World Endometriosis Society consensus on the classification of endometriosis. *Hum Reprod* 2017;**32**:315–324.

Jutel A. Classification, disease, and diagnosis. *Perspect Biol Med* 2011; **54**:189–205.

Keckstein J, Ulrich U, Possover M, Schweppe KW. ENZIAN-Klassifikation der tief infiltrierenden Endometriose. *Zentralbl Gynäkol* 2003;**125**:291–291.

Khazali S. Endometriosis classification—the quest for the Holy Grail? *J Reprod Infertil* 2016;**17**:67.

Kim JS, Lee CW, Yun J, Lee JH, Yun BH, Park JH, Seo SK, Cho S, Choi YS, Lee BS. Use of the endometriosis fertility index to predict natural pregnancy after endometriosis surgery: a single-center study. *Gynecol Obstet Invest* 2019;**84**:86–93.

Koninckx PR, Ussia A, Adamyan L, Wattiez A. An endometriosis classification, designed to be validated. *Gynecol Surg* 2011;**8**:1–6.

Leonardi M, Espada M, Vanza K, Choi S, Chou D, Chang T, Smith C, Rowan K, Condous G. Ultrasound is highly accurate at predicting the american society of reproductive medicine (AARM) stage of endometriosis. *J Minim Invasive Gynecol* 2019;**26**: S70–S71.

McCrorie AD, Donnelly C, McGlade KJ. Infographics: healthcare communication for the digital age. *Ulster Med J* 2016; **85**:71–75.

Montanari E, Dauser B, Keckstein J, Kirchner E, Nemeth Z, Hudelist G. Association between disease extent and pain symptoms in patients with deep infiltrating endometriosis. *Reprod Biomed Online* 2019;**39**:845–851.

Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertil Steril* 1997;**67**:817–821.

Rolla E. Endometriosis: advances and controversies in classification, pathogenesis, diagnosis, and treatment. *F1000Res* 2019;**8**:F1000 Faculty Rev-529.

Saridogan E, Byrne D. The british society for gynaecological endoscopy endometriosis centres project. *Gynecol Obstet Invest* 2013;**76**: 10–13.

Stiftung Endometriose Forschung. Consensus meeting, 15th Conference of the Stiftung Endometriose Forschung (Foundation for Endometriosis Research), Hotel Enzian, Weissensee, Austria, February 14–17, 2019. Weissensee, Austria: Stiftung Endometriose Forschung (SEF).

Timoh KN, Stewart Z, Benjoar M, Beldjord S, Ballester M, Bazot M, Thomassin-Naggara I, Darai E. Magnetic resonance enterography

to assess multifocal and multicentric bowel endometriosis. *J Minim Invasive Gynecol* 2018;**25**:697–705.

Tomassetti C, Geysenbergh B, Meuleman C, Timmerman D, Fieuws S, D'Hooghe T. External validation of the endometriosis fertility index (EFI) staging system for predicting non-ART pregnancy after endometriosis surgery. *Hum Reprod* 2013;**28**: 1280–1288.

Vanhie A, Meuleman C, Tomassetti C, Timmerman D, D'Hoore A, Wolthuis A, Van Cleynenbreugel B, Dancet E, Van den Broeck U, Tsaltas J *et al.* Consensus on recording deep endometriosis surgery: the CORDES statement. *Hum Reprod* 2016;**31**:1219–1223.

Vercellini P, Trespidi L, De Giorgi O, Cortesi I, Parazzini F, Crosignani PG. Endometriosis and pelvic pain: relation to disease stage and localization. *Fertil Steril* 1996;**65**:299–304.