ORIGINAL ARTICLE

Enhancing clinical practice: the Endoscore app for automated surgical data capture and endometriosis scoring

Arrigo Fruscalzo[1], Georgia Theodorou[1], Ambrogio Pietro Londero[2,3], Benedetta Guani[1], Jean-Marc Ayoubi [4], Anis Feki[1], Carolin Marti[1]

[1]Department of Gynaecology and Obstetrics, University Hospital of Fribourg, Fribourg, Switzerland
[2]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Infant Health, University of Genoa, Genoa, Italy
[3]Obstetrics and Gynaecology Unit, IRCCS Istituto Giannina Gaslini, Genova, Italy
[4]Department of Obstetrics and Gynaecology and Reproductive Medicine, Hopital Foch–Faculté de Médecine Paris, Suresnes, France

ABSTRACT

Background: There is a growing unmet need to digitalise the management of clinical data in medicine. Web-based scoring applications for endometriosis align with this trend.

Objectives: This study aimed to evaluate a web-based application that automatically calculates endometriosis staging scores [revised American Society for Reproductive Medicine classification (r-ASRM), the revised Enzian classification (#Enzian), Endometriosis Fertility Index (EFI)] and compare it to manual scoring in a proof-of-concept study.

Methods: 20 endometriosis cases operated on in 2022 were retrospectively selected. Six experienced gynaecologists were randomly allocated to either the conventional paper-based method or the digital application for staging of disease.

Main Outcome Measures: Completion time, score consistency among examiners and methods, and user satisfaction were recorded using a Likert scale and a subjective mental effort questionnaire (SMEQ).

Results: In comparison to the paper-based method, the web-based tool reduced scoring time by 25.1 seconds (128.0 vs. 153.1, $P<0.05$), was perceived as easier to use (higher Likert scale scores), and was associated with low-to-moderate mental effort on the SMEQ. The agreement between electronic and paper forms was rated as very good to excellent for r-ASRM [intraclass correlation coefficient (ICC): 0.93] and #Enzian (ICC: 0.84), while it was moderate for EFI (ICC: 0.67). Interrater agreement utilising the electronic form demonstrated high levels, yielding very good to excellent results for r-ASRM (ICC: 0.93) and EFI (ICC: 0.82) while showing moderate agreement for #Enzian (ICC: 0.63).

Conclusions: The application facilitates sequential data entry for users and automatically calculates r-ASRM, #Enzian, and EFI scores. It decreases scoring duration, strongly aligns with the paper-based method, and enhances user satisfaction.

What is New? This tool can potentially improve clinical efficiency, accuracy, and consistency in the staging of endometriosis.

Keywords: Endometriosis, classification, documentation, application, EFI, rASRM, #Enzian

Introduction

Endometriosis is characterised by the growth, adhesion, and progression of endometrial glands and stroma outside the uterine cavity, and it has significant implications for women of childbearing age.[1] While it affects up to 10% of the general female population, its prevalence can be as high as 50% among infertile women.[2] This chronic inflammatory disease relies on oestrogen for the establishment and proliferation of endometriotic tissue. Although endometriosis often appears benign on histopathologic examination,

it behaves like a malignant tumour, as lesions exhibit growth, infiltration, and adhesion to adjacent tissues, disrupting physiological processes.[3] Predominantly occurring in women of reproductive age, endometriosis is closely linked with infertility, chronic pelvic pain, and pelvic organ dysfunction.[4]

Much research has focused on defining and classifying the extent of endometriosis. However, there is a poor correlation between clinical symptoms and the extent of endometriotic lesions. The World Endometriosis Society recommends that all women with an endometriosis diagnosis undergoing surgery be staged according to the revised scoring system of the American Society for Reproductive Medicine r-ASRM[5,6]. Additionally, those with deep endometriosis should be staged according to the Enzian classification (#Enzian).[7] Finally, for women concerned about infertility, surgeons should calculate the Endometriosis Fertility Index (EFI) score to estimate their future fertility potential.[8]

Manual paper documentation is time-consuming and lacks automatic saving and cross-checking capabilities. Automatic electronic recording linked to a database is of essential importance in these cases of complicated multiple co-existing scores, especially since they can be traced back easily for statistical and research purposes, as well as patient follow-up optimisation.[9] The accurate documentation of the disease extent and localisation during surgery is one of the core issues of the treatment process of endometriosis for future study purposes.

In this proof-of-concept study, we aimed to evaluate the performance of a web-based application that automatically calculates endometriosis staging scores according to the three currently most used classifications of endometriosis, the r-ASRM, the revised #Enzian classification (#Enzian) and EFI. We also compared the use of this application to classical manual scoring and evaluated its users' acceptance in daily practice.

Methods

The Fribourg Cantonal Hospital (HFR) commissioned an external company to develop the application. Several meetings between the gynaecology and obstetrics department's physicians and the company helped define the required features and user interface as accurately as possible. For this purpose, the core issues of each score were selected and evaluated. Information concerning different scores, if present, was used to calculate multiple scores at a glance. These issues were extrapolated by analysing the original manuscripts presenting the three scores concerned.

Data collected to enable classification and calculation of the three scores were anamnestic and surgical. Anamnestic patients' data included the birth date, which was used to calculate the patient's age, previous pregnancy history, and infertility history of at least 3 years, if present, necessary to calculate the EFI score.

Surgical data included recording the presence of endometriosis and the extension of lesions in terms of larger diameter and total area at the following sites:

• Peritoneum-the sum of the area in cm

• Posterior cul-de-sac-partial or complete obliteration, left and right ovary, the sum of the diameter in cm, respectively, superficial or deep

• Ovarian and fallopian tubes (left and right) - enclosure through adhesions, respectively, filmy or dense

• Fallopian tube patency (left and right)-positive or negative, if tested

• Rectovaginal space or retro-cervical area-in cm

• Utero-sacral ligaments (left and right)-in cm

• Recto-sigmoid-in cm

• Least adnexal function (left and right) according to the EFI score

• Further localisations-presence of adenomyosis; infiltration of the bladder, the ureters, intestine (other than recto-sigmoid)

We then conducted a proof-of-concept study concerning the use of the application compared to the classic analogic paper-based classification and scoring system. For this scope, we retrospectively assessed endometriosis scoring and classification using the two methods, which were performed on a subset of patients randomly chosen from those diagnosed with endometriosis and operated on in 2022. A group of six surgeons from the service of gynaecology at the HFR, with at least two years of experience in the field of endometriosis, were selected. They were randomly assigned to either the classic analogic classification or the digital group.

The selected surgeons examined the operation protocols of the 20 patients and reviewed operation images and histological reports. They were then asked to describe and classify the surgical status by completing the paper format (Appendix 1) or by the digital format, respectively.

They were then asked to perform a subjective evaluation of the application through a brief questionnaire examining accessibility, ease of use, and estimation of clinical impact on daily practice using a Likert-type scale and the Subjective Mental Effort Questionnaire (SMEQ) questionnaire (Appendix 2).[10] The time required to calculate the scores and the robustness in score calculation between examiners and between the two methods were also evaluated. Finally, the intraclass correlation coefficients (ICC) were calculated to assess the measurement's reliability. Usually, it is suggested that ICC values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability and values greater than 0.90 indicate excellent reliability.[11]

We calculated a sample size of 16 patients to have an 80% chance of detecting, at a significant level of 5%, a decrease of one standard deviation in the primary outcome measure (time estimated for calculating the three scores). For instance, this could entail a reduction from 5 minutes in the control group (paper-based) to 3 minutes in the experimental group (application-based), assuming a standard deviation of the difference of 2 minutes. An additional 4 patients were included, bringing the total to 20, considering a 20% possibility of missing or non-interpretable results.

Statistical Analysis

Statistical analysis was performed utilising R Studio version 4.4.2. Data were depicted as mean values accompanied by their respective standard deviations or as medians with interquartile ranges. For datasets conforming to a normal distribution, the Student's t-test was employed. The administration of the SMEQ questionnaire was carried out among experts, with median responses computed alongside their interquartile range. A comparative analysis of paper-based versus electronic compilation duration was conducted, with mean values and standard deviations being presented. The concordance between paper and electronic evaluations was evaluated utilising the mean together with a 95% confidence interval (CI), as was the inter-operator concordance for electronic assessments.

This study was accepted by the Cantonal Ethics Committee of Vaud (Commission cantonale d'éthique de la recherche sur l'être humain, CER-VD), Lausanne, Switzerland (CER-VD number 2021-00258). Informed consent (general consent of HFR) was obtained for each patient included in the study.

Results

From a technical standpoint, the application was built using the Next.js framework (based on React), which facilitates the creation of modern, high-performance web applications. The application is responsive and designed to function across various devices, including computers, tablets, and smartphones. It is hosted within HFR and is accessible at https://app.h-fr.ch/endo.

Regarding the user interface, the application uses a stepper system to guide users through the step-by-step entry of patient data. The extent of endometriosis concerning different localisations is entered. Once the data is entered, the software automatically calculates the r-ASRM, #Enzian, and EFI scores at a glance. For this scope, a Logic arbour was generated (Figure 1). This stepper-based approach ensures easy and structured navigation, minimising errors and enhancing the user experience (Figure 2). The entered data is temporarily stored in the browser, allowing for modifications if needed. These data are finally displayed as a summary that can be exported in PDF format.

The Likert scale results provided by the participants showed an advantage in the use of the web-based application in terms of rendering everyday practice easier since navigating and using the interface and the infographic scoring of the application were easier overall than the paper format score (Figure 3a). The SMEQ confirms these results, showing that the use of the web-based application was of low-moderate difficulty for the participants. No members of the surgical panel reported increased difficulty in the use, whereas two rated scores 20 or less on the SMEQ chart, with the median set at 35 (Figure 3b). Regarding time comparison between the two evaluated methods, the web application proved less time-consuming by 25.1 seconds (electronic 128.0 sec (±56.7) vs. paper 153.1 (±73.7), $P<0.05$) (Figure 4a).

Concerning the robustness of the rASRM, #Enzian, and EFI scores calculation, we evaluated the agreement between the electronic and paper forms. Accordingly, to the current interpretation of the ICC values, rASRM, EFI, and #Enzian scores showed very good and excellent results, respectively, with an ICC of 0.93 (95% CI: 0.48-
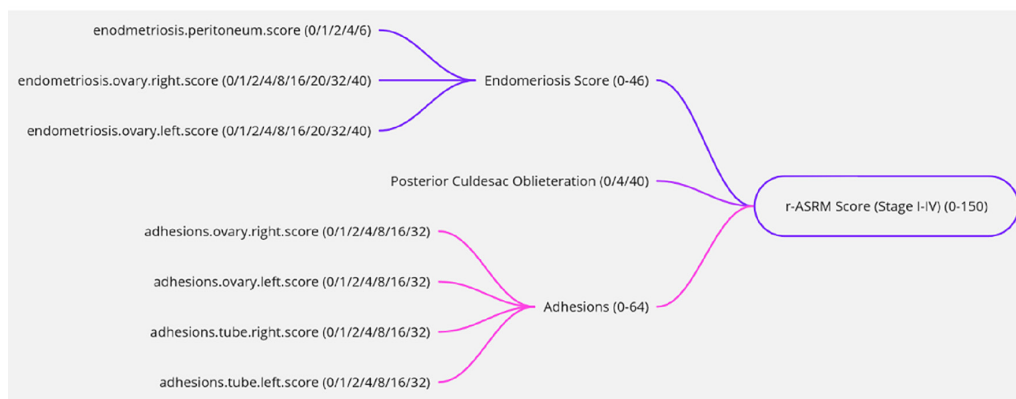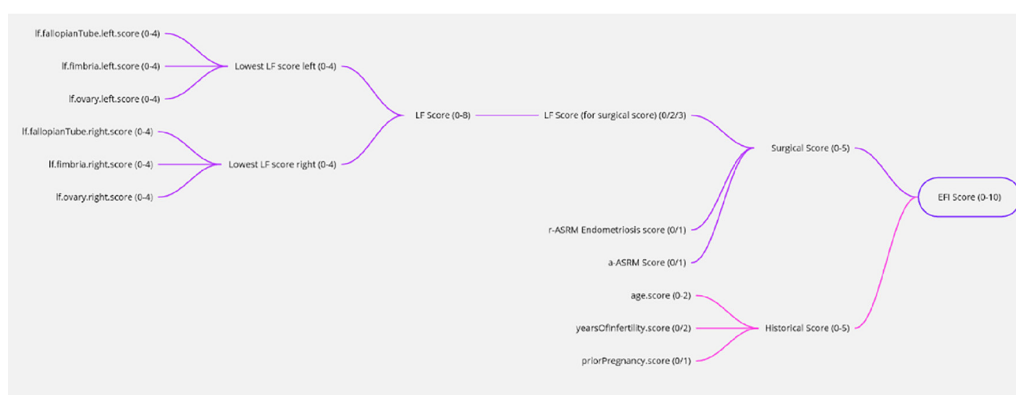
Figure 1. Logic arbour generated to calculate the r-ASRM (Panel A) and the EFI (Panel B) scores.

r-ASRM: Revised American Society for Reproductive Medicine classification, EFI; Endometriosis fertility index.

0.98) for rASRM of 0.96 (95% CI: 0.90-0.98) for EFI, and of 0.84 (95% CI: 0.79-0.88) for the #Enzian score (Figure 4b).

Regarding the agreement between users using the electronic form, the r-ASRM and EFI scores showed very good and excellent results, respectively, with an ICC of 0.93 (95% CI: 0.48-0.98) for r-ASRM and of 0.82 (95% CI: 0.59-0.93) for the EFI score. The #Enzian score showed a moderate concordance, with an ICC value of 0.63 (95% CI: 0.56-0.69) (Figure 4c).

Discussion

Main Findings

There is a growing unmet need to digitalise and, therefore, simplify the management of clinical data in medicine. We successfully developed and tested the feasibility of using a progressive web-based application for endometriosis staging and classification. The comparison of results between analogic and digital evaluations demonstrated high reliability across all three scores: the rASRM, the EFI,

and the #Enzian scores. The strong interrater agreement observed in the results for the r-ASRM and EFI scores, and to a lesser extent for the #Enzian score, validates the reliability of the scoring outcomes generated by the web-based application.
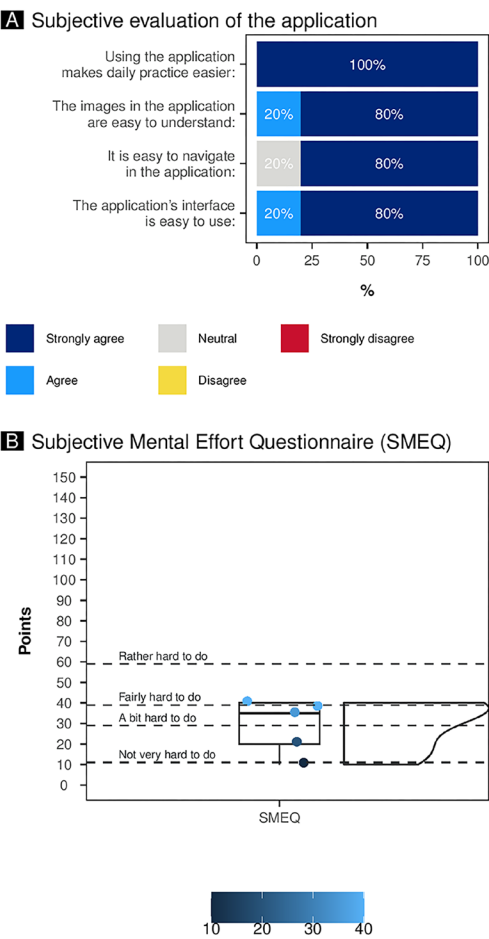
Strengths and Limitations

The strengths of this study were the inclusion of six gynaecologists of various degrees of experience, and agreement was found to be unanimously acceptable on their satisfaction and the time saved with the web application. Furthermore, there is the potential to develop further versions of the application, for example, describing the pathology's preoperative extent at the clinical examination or during imaging (ultrasound, magnetic resonance imaging, or others).

Of interest, we tested the application during a national congress of gynaecology and obstetrics. Forty-nine gynaecologists who participated in the congress were asked to test the application on a fictitious case of

Figure 2. Screenshots of the web-based application user's interface.

Screenshots of the web-based application user's interface for the step-by-step data entry according to localisation and extent of endometriosis lesions and adhesion: 1) Patient's anamnestic data; 2) Peritoneum; 3) Left ovary (right ovary not shown); 4) Posterior cul-de-sac; 5-6) Left tube (right tube not showed); 7-10) Deep infiltration endometriosis of the rectovaginal space and retrocervical area, left sacrouterine ligament and left pelvic sidewall (right sacrouterine ligament and right pelvic sidewall not shown), rectum, as well as other localisations, according to the Enzian (#Enzian) classification; 11) Least function left (right side not shown) according to the EFI classification; 12) A summary of the three scores according to the r-ASRM, #Enzian and EFI classification.

rASRM: Revised American Society for Reproductive Medicine classification, #Enzian: Revised Enzian classification, RFI: Endometriosis Fertility index.

A Subjective evaluation of the application

Using the application makes daily practice easier: 100%

The images in the application are easy to understand: 20% 80%

It is easy to navigate in the application: 20% 80%

The application's interface is easy to use: 20% 80%

%

Strongly agree · Neutral · Strongly disagree · Agree · Disagree

B Subjective Mental Effort Questionnaire (SMEQ)

Points

Rather hard to do

Fairly hard to do
A bit hard to do

Not very hard to do

SMEQ

10 20 30 40

Figure 3. Evaluation of the electronic application utility using: Panel A) a Likert-type scale; Panel B) the Subjective Mental Effort Questionnaire questionnaire.

endometriosis. All were able to complete the scoring via the application within the allocated three minutes, demonstrating excellent compliance and providing very positive feedback regarding its ease of use (data not shown). Even though the extent of time saved in the calculation of the scores ranged only in terms of a few seconds, it was important to test that the application was at least not more time-consuming than the common paper-based form. In addition, although showing little clinical relevance in isolation, this saving can add to significant efficiency gains in high-volume centres or multi-case research settings.

One current limitation of our application is that clinical notes are not saved in a digital format, but only summarised and stored in a PDF format. This is to reduce the constraints linked to personal data management regulation encountered when digital data is stored. An upgraded version of the application will allow the automated integration of data in a patient database. This will also be greatly useful for internal quality checks and research purposes. Another limitation of our results is the retrospective design of the proof-of-concept study. Also, the relatively small number of unselected patients included cannot permit generalisation of the results we presented. A prospective external evaluation and a prospective multicentric study should be performed to further test the clinical relevance of our results in everyday practice.

Strengths and Limitations Compared to Other Studies

Metzemaekers et al.[12] proposed the use of an application in alignment with these efforts to provide uniform surgical registration and classification of endometriosis. The Endometriosis Quality and Grading Instrument for Surgical Performance was the first recorded attempt that used infographics and presented an applicable digital scoring system with immediate use in clinical practice. The study was performed on a group of experts in the
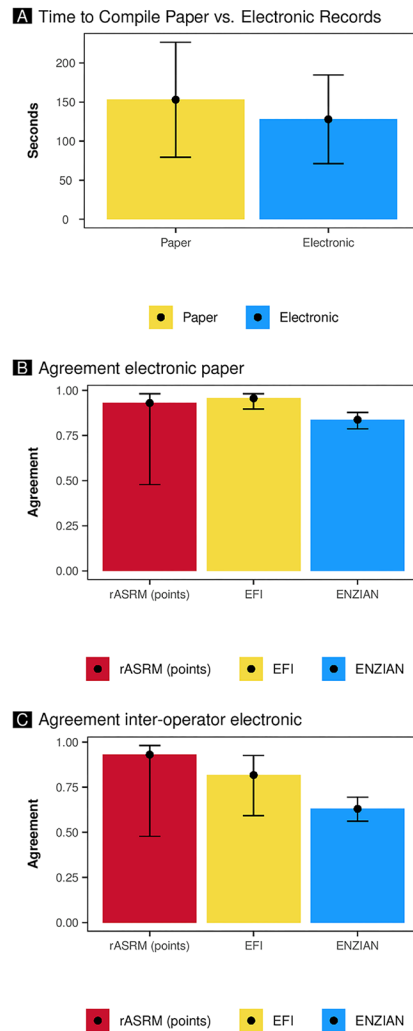
Figure 4. Completion time for calculating endometriosis staging scores using the r-ASRM, #Enzian and EFI classification systems. Panel A) Average time (± standard deviation) needed to compile case records using paper compared to electronic methods; Panel B) Agreement (95% CI) between electronic and paper scoring systems; (Panel C) Inter-operator agreement (95% CI) for electronic scoring of each classification system.

rASRM: Revised American Society for Reproductive Medicine classification, #Enzian: Revised Enzian classification, RFI: Endometriosis Fertility index, CI: Confidence interval.

field of endometriosis, randomly assigned to stage a fictive case of severe endometriosis either in a classical way or using web-based software. The result showed that the digital application was superior in the correct evaluation of the r-ASRM numerical score and stage, as well as the #Enzian score. Also, the application's usability, measured with the system usability scale and the SMEQ, was found to favour the digital evaluation.[12] However, this application does not allow the digital calculation of the revised version of the #Enzian and of the EFI score, as our application does.

The application we presented is specifically designed to describe the localisation and extent of endometriosis during and out of surgery. This application is accessible via a browser and available on personal computers and smartphones. It thoroughly guides the gynaecologist step by step in recording the surgical localisation and extent of endometriosis at surgery. The results generate a PDF sheet documenting the scores according to the three most used classification systems: the r-ASRM,, #Enzian, and EFI scores. The file can be added to the patient's record for storage purposes.

The web-based system was selected for various advantages. The system facilitates the concurrent description and computation of the three endometriosis scores through an intuitive slider-based interface. The

solution guarantees total independence from particular platforms and architectures, facilitating straightforward deployment as a Docker container on either local or web farm servers. The platform is compatible with various devices, including x86-based computers, ARM-based computers, smartphones and tablets, Apple Silicon laptops, and emerging RISC-V devices. The application is accessible via any modern browser, thereby negating the necessity for installations, addressing dependency conflicts, and circumventing multiple builds or compatibility layers.

Our results indicate that the slider-based system facilitated the simultaneous description and calculation of the three scores, unlike the traditional paper-based system, which requires separate processing for each classification system. The parameter-driven design utilising structured logic trees facilitates integration with LLM-based decision support systems and other automated pipelines. The digital endometriosis scoring and classification system operates more efficiently and rapidly than the traditional paper-based documentation system. This finding is consistent with existing literature indicating that digital medical documentation and scoring systems provide multiple benefits compared to traditional paper-based methods.[13,14]

Clinical and Policy Implications

The accurate documentation of the disease extent and localisation during surgery is one of the core issues of the clinical management process of endometriosis and is essential for collecting data for study purposes. The documentation of the surgical status according to the three aforementioned classification systems, r-ASRM,, #Enzian, and the EFI has become an integral part of many centres that treat endometriosis, aiming to centralise and uniform staging of this disease.

We have provided an easy-to-use and practical web application for registering localisation and extent of endometriosis, providing at a glance the score of the three different classifications currently mostly used, rASRM, #Enzian, and the EFI. Providing a user-friendly and fast application for this purpose is crucial for hospitals treating complex conditions like endometriosis, as these imply saving time for specialised personnel and standardised patient data storage.

Unanswered Questions and Future Research

The use of digital applications in medicine is increasing but remains a novelty requiring research and development as well as suitable training to allow for their integration into daily practice. Several apps have been developed to evaluate the clinical symptoms of endometriosis.[15] According to this review, there are currently 12 applications used for this scope. Applications have been shown to substantially help patients and physicians in this process, and several international organisations support their use. Nonetheless, this review highlights the scarcity of effective patient and physician-directed apps for managing and treating endometriosis, particularly for the intraoperative description and management of surgical data. This argument demonstrates the considerable potential for developing mobile resources to aid in diagnosing and treating endometriosis.
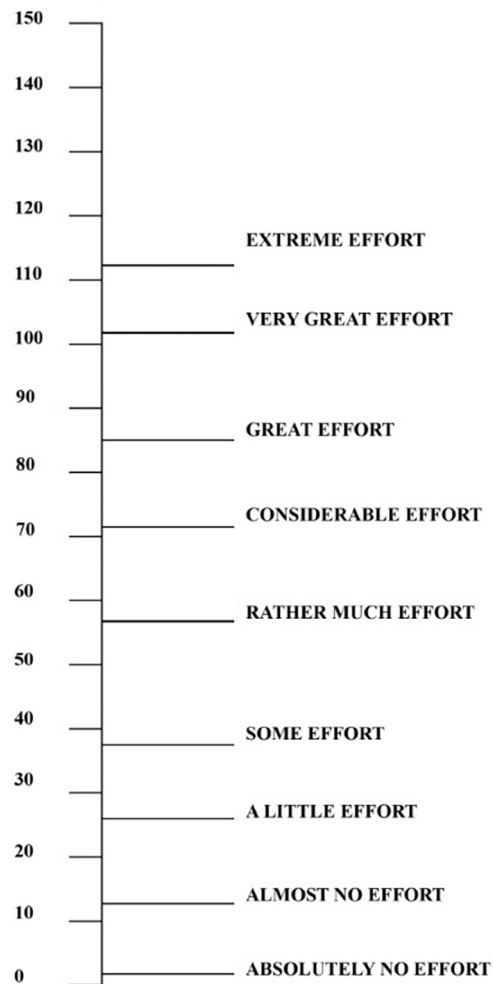
References

1. Bloski T, Pierson R. Endometriosis and Chronic pelvic pain: unraveling the mystery behind this complex condition Nurs Womens Health. 2008;12:382-95.

2. Missmer SA, Hankinson SE, Spiegelman D, Barbieri RL, Marshall LM, Hunter DJ. Incidence of laparoscopically confirmed endometriosis

by demographic, anthropometric, and lifestyle factors. Am J Epidemiol. 2004;160:784-96.

3. Kobayashi H, Imanaka S, Yoshimoto C, Matsubara S, Shigetomi H. Rethinking the pathogenesis of endometriosis: Complex interactions of genomic, epigenetic, and environmental factors. J Obstet Gynaecol Res. 2024;50:1771-84.

4. Bonavina G, Taylor HS. Endometriosis-associated infertility: From pathophysiology to tailored treatment. Front Endocrinol (Lausanne). 2022;13:1020827.

5. Johnson NP, Hummelshoj L; World Endometriosis Society Montpellier Consortium. Consensus on current management of endometriosis. Hum Reprod. 2013;28:1552-68.

6. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. Fertil Steril. 1997;67:817-21.

7. Keckstein J, Saridogan E, Ulrich UA, Sillem M, Oppelt P, Schweppe KW, et al. The #Enzian classification: A comprehensive non-invasive and surgical description system for endometriosis. Acta Obstet Gynecol Scand. 2021;100:1165-75.

8. Adamson GD, Pasta DJ. Endometriosis fertility index: the new, validated endometriosis staging system. Fertil Steril. 2010;94:1609-15.

9. Stujenske TM, Mu Q, Pérez Capotosto M, Bouchard TP. Survey analysis of quantitative and qualitative menstrual cycle tracking technologies. Medicina. 2023;59:1509.

10. Bevan N. Measuring usability as quality of use. Software Qual J. 1995;4:115–30.

11. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability Rresearch. J Chiropr Med. 2016;15:155-63.

12. Metzemaekers J, Haazebroek P, Smeets MJGH, English J, Blikkendaal MD, Twijnstra ARH, et al. 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. Hum Reprod Open. 2020;2020:hoaa053.

13. Stausberg J, Koch D, Ingenerf J, Betzler M. Comparing paper-based with electronic patient records: lessons learned during a study on diagnosis and procedure codes. J Am Med Inform Assoc. 2003;10:470-7.

14. Wu CHK, Luk SMH, Holder RL, Rodrigues Z, Ahmed F, Murdoch I. How do paper and electronic records compare for completeness? A three centre study. Eye (Lond). 2018;32:1232-6.

15. Baxter B, Khalil S, Seaman C, Seckin S, Ascher-Walsh C, Zakashanksy K. Rating endometriosis apps using the applications scoring system. Am J Obstet Gynecol. 2021;224:S801.

Appendix 1. Assessment of accessibility, usability, and clinical impact using likert scale and SMEQ.

	Statement	5 – Strongly agree	4 – Agree	3 – Neutral	2 – Disagree	1 – Strongly disagree
1	The application's interface is easy to use.					
2	It is easy to navigate through the application.					
3	The images within the application are easy to understand.					
4	Compared to the traditional paper-based method, the application facilitates daily documentation and staging practice.					

```
150 ——
140 ——
130 ——
120 ——
                    EXTREME EFFORT
110 ——
                    VERY GREAT EFFORT
100 ——
90 ——
                    GREAT EFFORT
80 ——
                    CONSIDERABLE EFFORT
70 ——
60 ——
                    RATHER MUCH EFFORT
50 ——
40 ——
                    SOME EFFORT
30 ——
                    A LITTLE EFFORT
20 ——
                    ALMOST NO EFFORT
10 ——
0 ——                ABSOLUTELY NO EFFORT
```

Appendix 2. The subjective mental effort question (SMEQ).

It is a single-item questionnaire with a rating scale from 0 to 150 with nine verbal labels ranging from "Not at all hard to do" (just above 0) to "Tremendously hard to do" (just above 110).