

A new validated screening method for endometriosis diagnosis based on patient questionnaires

Charles Chapron,^{a,b,c,1*} Marie-Christine Lafay-Pillet,^{b,1} Pietro Santulli,^{a,b,c} Mathilde Bourdon,^{a,b} Chloé Maignien,^b Antoine Gaudet-Chardonnet,^b Lorraine Maitrot-Mantelet,^b Bruno Borghese,^{a,b,c} and Louis Marcellin,^{a,b,c}

^aUniversité de Paris, Faculté de Santé, Faculté de Médecine Paris Centre, Paris, France

^bAssistance Publique – Hôpitaux de Paris (AP-HP), Hôpital Universitaire Paris Centre (HUPC), Centre Hospitalier Universitaire (CHU) Cochin, Department of Gynecology Obstetrics II and Reproductive Medicine (Professor Chapron), Paris, France

^cDepartment “Development, Reproduction and Cancer”, Institut Cochin, INSERM U1016, Paris, France

Summary

Background The time between symptoms onset and endometriosis diagnosis is usually long. The negative impacts of delayed endometriosis diagnosis can affect patients and health outcomes.

Methods We conducted a case-control study using clinical symptoms and epidemiological data extracted from a prospective pre-operative patient questionnaire compared between patients with histologically proven endometriosis and patients with no endometriosis at surgical exploration from 2005 to 2018, in a French referral center. We used the beta coefficients of the significant variables introduced in a multiple regression model to devise a score (score 1), evaluated by the area under the curve (or C-index), with three levels, defined by a score between 1 and ≥ 25 : (i) highly specific, identifying correctly the patients without the disease; (ii) highly sensitive, identifying the patients with the disease; and (iii) a level maximizing sensitivity and specificity for the best classification of the whole population. To minimize patient self-evaluation of pain, we devised a second score (score 2) with the same method and levels and scores definition, excluding visual analog scale pain scores, except for dysmenorrhea. These scores were validated on an internal and external population.

Findings Score 1 had a C-index of 0.81 (95% CI [0.79–0.83]). Results for the three score 1 levels were: ≥ 25 : specificity of 91% (95% CI [89–93]); < 11 : sensitivity of 91% (95% CI [89–93]); ≥ 18 : specificity of 75% (95% CI [72–78]) and sensitivity of 73% (95% CI [70–76]). Score 2 had a C-index of 0.75 (95% CI [73–77]). The three levels of score 2 were: ≥ 24 : specificity of 82% (95% CI [80–85]); < 7 : sensitivity of 92% (95% CI [90–94]); ≥ 17 : specificity of 62% (95% CI [58–65]) and sensitivity of 78% (95% CI [75–81]). The two scores were internally and externally validated.

Interpretation A score based only on a patient questionnaire could allow identification of a population at high risk of endometriosis. This strategy might help referral to specialized radiologists for a non-surgical endometriosis scan.

Funding None.

Copyright © 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Endometriosis; Clinical diagnosis; Questionnaire; Imaging work-up; New paradigm; Clinical score; Multivariate regression analysis; External validation

Introduction

Endometriosis, defined as the presence of endometrial-like tissue outside the uterine cavity, is responsible for pain and infertility, and it hence has a major impact on quality of life.¹ Endometriosis is a heterogeneous, chronic disease with a systemic impact,² with three

clinical phenotypes: superficial peritoneal endometriosis (SPE), ovarian endometrioma (OMA), and deep infiltrating endometriosis (DIE).

Throughout the world, and irrespective of the health care system,³ the delay between the onset of symptoms and endometriosis diagnosis is between 6 and 10

*Corresponding author at: Department of Gynecology Obstetrics II and Reproductive Medicine, CHU Cochin, Bâtiment Port Royal, 123 boulevard Port Royal, 75014 Paris, France.

E-mail address: charles.chapron@aphp.fr (C. Chapron).

¹ Charles Chapron and Marie-Christine Lafay-Pillet contributed equally to this work and should be considered as joint first authors.

Research in context

Evidence before this study

Numerous clinical symptoms and epidemiological factors can be involved in the diagnosis of endometriosis. We searched PubMed for articles from 2000 to 2020 that documented the clinical predictive symptoms and the epidemiological data that can help achieve an early prediction of endometriosis, using the search terms “endometriosis”, “diagnosis”, and “questionnaire”. There are many congruent studies in terms of clinical symptoms related to endometriosis, essentially pain symptoms but also reproductive characteristics (infertility, low gravidity, and parity), epidemiological factors such as a family history of endometriosis. On the other hand, there is no consensus for other factors such as body mass index, menstrual cycle characteristics, age at menarche, birthweight, and blood group. Although some data support the fact that the localization and severity of endometriotic lesions may be predictive for the diagnosis, the reproducibility has not been evaluated. Few studies to date have combined all of the clinical factors to devise a predictive model with high performances, and none have employed an external validation.

Added value of this study

This study used all of the available variables selected from our prospective database introduced in a multivariate regression analysis to improve the prediction of endometriosis diagnosis. We generated a clinical score in order to obtain better performances for selecting patients at high versus low risk of the disease, and the results were validated on an internal and an external population, which has not been done before.

Implications of all the available evidence

As it generally takes a long time for endometriosis to be diagnosed, earlier diagnosis with a simple score based on a patient questionnaire could help provide timely treatment of young patients, thereby avoiding severe complications and improving quality of life. Such a score would allow the time involved in reaching a diagnosis to be shortened and it would improve the management (medical treatment, surgery, assisted reproductive technologies), the wellbeing, and the quality of life of patients.

years.⁴ This delay has major consequences not only for the patient (infertility, decrease of the ovarian reserve, major complications [intestinal occlusion, loss of a kidney, etc.], sexuality, relationship with their partner, fatigue, depression, central sensitization with consequences at the brain and/or peripheral levels, degradation of the patient-practitioner relationship, etc.) but also for society as a whole (loss of work productivity and a substantial economic burden).¹ Initially considered mainly in terms of its impact on the afflicted

individuals, endometriosis is increasingly recognized as a major public healthcare and societal problem,⁵ with a prevalence of approximately 10% of the female population.^{1,2}

These reasons warrant revisiting endometriosis diagnostic modalities.¹ The “revolution” in the way endometriosis is diagnosed was made possible by the substantial improvements in gynecological imaging in recent years. Indeed, transvaginal ultrasonography (TVUS) and/or magnetic resonance imaging (MRI) are used nowadays to diagnose endometriosis, to establish its phenotype (OMA and DIE), to determine the exact localizations of the lesions,⁶ as well as to assess the association with adenomyosis⁷ with excellent predictive performance. This is the reason why several gynecological scientific societies are in agreement that patients who are experiencing pain from endometriosis and who do not have an immediate desire for pregnancy can be prescribed hormonal therapy without histological proof of endometriosis.⁸⁻¹²

However, referring all patients suspected of having endometriosis to specialized radiologists is not feasible. An initial screening to select the patients at higher risk of the disease is needed. A patient questionnaire is the first step in the diagnostic process. When properly conducted, it is sufficient to detect patients at high risk of endometriosis who should be referred to a specialized radiologist who is aware of the specific imaging modalities for the disease.¹

Therefore, the aim of this study was to devise and validate a score based solely on a patient questionnaire to select a sample of patients with a high prevalence of endometriosis using cut-off points with high specificity and sensitivity.

Methods

We used a single-center, prospective cohort to devise a predictive score for endometriosis based on clinical variables from a simple questionnaire. This score was validated on an internal and an external population.

Study design and the participants

We performed a case-control study with an internal and an external validation sample.

We used a prospective database of symptomatic < 42-year-old patients surgically explored in a teaching hospital for benign gynecological pathologies from 2005 to 2018. Pregnant patients and patients operated on for cancer were not included, nor were patients who refused to provide consent.

The patients included for the analysis were divided into two groups: (i) a control group comprising patients without visual endometriosis lesions, as verified during the surgical procedure; and (ii) a study group comprising women with histologically proven endometriotic

lesions after complete excision.¹³ Because endometriotic phenotypes are frequently associated, patients were classified according to their worst lesion, from least to most severe, as follows: SPE, OMA, and DIE.¹³

One-third of the population was randomly extracted for an internal validation and two-thirds were used as the derivation cohort.¹³ An external validation was carried out on an incident case-control study on a Russian population using the same prospective data collection, inclusion criteria,¹⁴ and endometriosis phenotype classification.¹² For every DIE patient recruited, two women without endometriosis, two with SPE, and two with OMA were prospectively recruited between May 2011 and April 2013.¹⁴

The study and the establishment of the database were approved by the local institutional review board (approval number 05-2006 issued by the “Comité de Protection des Personnes et des Biens dans la Recherche Biomédicale” of the Cochin Hospital, Paris, France). Written consent was obtained from all included patients. For the external validation sample, the investigator/institution obtained approval from the Independent Ethics Committee/Institutional Review Board as applicable in the country of the study.¹⁴ All the data were fully anonymized before use.

Data collection

The data were prospectively collected by the surgeon during a face-to-face interview to complete a patient questionnaire in the month before the surgery.¹³ Five painful symptoms (primary or secondary dysmenorrhea, dyspareunia, pain of gastrointestinal (GI) origin, pain from urinary tract symptoms, and non-cyclical chronic pelvic pain) were evaluated using visual analog scales (VAS) from 0 to 10.¹³ Primary dysmenorrhea was defined as painful menstruations occurring shortly after menarche, although it could also occur as late as one year after menarche.¹⁵ Secondary dysmenorrhea was defined as painful menstruations occurring at a later time. The GI symptoms were defined as pain when defecating at the time of menstruation or intestinal cycle pain.¹³ Lower urinary tract (LUT) symptoms were defined as one or more of the following symptoms during menstruation: hematuria, recurrent urinary tract infections, pain when urinating, dysuria, and non-microbial cystitis. Chronic pelvic pain was defined as non-cyclical pain located in the pelvic area, other than dyspareunia.¹³ In the course of completing the questionnaire, the patients were asked about their worst pain in the past 6 months.¹³ If the patient had no menstruation for 6 months, the dysmenorrhea VAS score was 0, as well as the other variables defined by cyclic variations (GI and LUT symptoms). If the patient had no sexual activity, the dyspareunia VAS score was 0.¹³

We collected data on the characteristics of the menstrual cycles, family history of endometriosis, and parity

and gravidity. Infertility was defined as at least 12 months of unprotected intercourse that did not result in pregnancy. Patients were deemed to be non-infertile if they were not trying to become pregnant.

Statistical analysis

The combination of macroscopic and histological data served as the gold standard for definition of the lesions and patient classifications as SPE, OMA, or DIE.¹³ Based on a comprehensive literature review and a clinical rationale, we identified 34 variables of the pre-operative questionnaire data previously entered in the database and potentially associated with endometriosis.

Student's *t*-test was used for the continuous data and the chi-square test was used for the categorical data to select the variables of the questionnaire associated with the presence of endometriosis at a threshold of $p \leq 0.30$. The continuous variables were then converted into dichotomous variables using receiver operating characteristic (ROC) curves. Composite variables were created in case of common clinical significance or eliminated based on clinical criteria considered useless for early clinical prediction (such as prior surgery for endometriosis) or potentially inaccurate (such as menorrhagia).

All of the selected variables were introduced in a multiple logistic regression model to eliminate any interaction and to select, by a backward stepwise procedure, the best combination of variables independently associated with the diagnosis of endometriosis ($p \leq 0.05$). Adjusted diagnostic odds ratios (aOR) were calculated with their 95% confidence intervals (95% CI). A Jackknife procedure was applied to the model to detect variables potentially responsible for instability.¹⁶ We then chose a simplified model by reduction of the number of variables that did not have a major influence on the performance of the model. The impact of excluding variables was evaluated on the R^2 of the multiple regression analysis and the ROC area of the score. The performance of the final model for predicting endometriosis was specified by the area under the curve (C-index). The calibration of the model was evaluated using the Hosmer-Lemeshow test.¹⁷

We devised a score using the β coefficients from the multivariate analysis.¹⁸ Calculation of the scores for each participant was performed as previously published¹²: We rounded up the coefficients from the multivariate analysis¹² and multiplied them by ten for easier use. We then multiplied this coefficient with the value of the corresponding variable (0 or 1) and we calculated the sum to obtain the value of the score for all patients.

We then defined three levels for the score according to their sensitivity, specificity, and positive and negative likelihood ratio: (i) a very high risk maximizing the specificity so as to exclude the maximum number of controls; (ii) a low-risk score favoring the sensitivity to

Variables	N = 885	N = 800	Controls N = 885 moyenne ± sem or n (%)	Patients N = 800 mean ± sem or n (%)	p
Age (year)	885	799	32.0 ± 1.19	31.5 ± 1.18	0.060 ^a
Height (cm)	882	797	164.9 ± 0.22	165.7 ± 0.22	0,020 ^a
Weight (kg)	882	797	63.0 ± 0.38	60.19 ± 0.37	< 0,001 ^a
Birth weight (g)	388	468	3214 ± 29.50	3158 ± 27.90	0,17 ^a
BMI (kg/m ²)	878	798	23.1 ± 0.10	21.9 ± 0.10	< 0,001 ^a
Gravidity (number of pregnancies)	883	800	1.00 ± 0.05	0.5 ± 0.03	< 0,001 ^a
Parity (number of deliveries)	883	800	0.50 ± 0.03	0.25 ± 0.03	< 0,001 ^a
Family history of endometriosis	885	800	19 (4,0)	87 (11,0)	<0,001 ^b
Sterility none primary secondary	885	796	592 (66,9) 159 (18,0) 134 (15,1)	512 (64,3) 210 (26,4) 74 (09,3)	<0,001 ^b
Duration of sterility (months)	264	264	39.3 ± 1,9	40.1 ± 1,9	0,740 ^a
Pain	885	800	364 (41.1)	606 (75.7)	<0,001 ^b
Duration of pain (months)	331	579	36.6 ± 2.6	56.3 ± 2.6	< 0,001 ^a
Dysmenorrhea no yes primary yes secondary	882	795	368 (41.7) 305 (34.6) 209 (23.7)	96 (12.1) 405 (50.9) 294 (37.0)	0,01 ^b
VAS pains scores					
VAS dysmenorrhea	880	794	4.1 ± 0.1	6.9 ± 0.1	< 0,001 ^a
VAS deep dyspareunia	844	770	2.2 ± 0.1	4.3 ± 0.1	< 0,001 ^a
VAS non cyclic pelvic pain	881	797	1.7 ± 0.1	3.2 ± 0.1	< 0,001 ^a
VAS GI symptoms	882	797	0.8 ± 0.1	3.7 ± 0.1	< 0,001 ^a
VAS urinary tract symptoms	878	796	0.2 ± 0.03	1.1 ± 0.1	< 0,002 ^a
Age at first menstruation	875	793	12,9 ± 0,06	12,9 ± 0,06	0,440
Absence from school because of primary dysmenorrhea	872	792	144 (16.5)	247 (31.2)	<0,001 ^b
Loss of consciousness due to dysmenorrhea	887	800	37 (4.2)	108 (13.5)	<0,001 ^b
OCP for severe primary dysmenorrhea	305	405	72 (23.6)	180 (44.4)	<0,001 ^b
Rectorrhagia	883	795	28 (3.2)	112 (14.1)	<0,001 ^b
Hematuria	881	792	18 (2.0)	39 (4.9)	<0,001 ^b
Menorrhagia	873	789	386 (44.2)	380 (48.2)	0,002 ^b
Length of menstruation (days)	811	759	5.2 ± 0,08	5,3 ± 0,07	0,688 ^a
OCP never ever current	880	793	586 (66.6) 81 (9.2) 213 (24.2)	458 (57.8) 103 (13.0) 232 (29.2)	0,001 ^b
prior surgery for endometriosis	882	800	34 (3.8)	270 (33.7)	< 0,001 ^b
Prior surgery for endometrioma	882	792	10 (1.1)	116 (14.6)	< 0,001 ^b
Negative rhesus blood group	481	472	66 (13.7)	84 (17.8)	0.05 ^b
Menstrual cycles regular often regular irregular	874	792	665 (76.1) 34 (3.9) 175 (20.0)	620 (78.3) 13 (1.6) 159 (0.2)	0.02 ^b
Duration of regular cycles	654	614	28.3 ± 0.17	28.2 ± 0.10	0.696 ^a
Age at first pregnancy	220	130	25.5 ± 0.39	27.3 ± 0.37	0.002 ^a
Age at first OCP	408	520	19.9 ± 0.23	18.2 ± 0.14	< 0,001 ^a

Table 1: bivariate analysis of variables comparing controls and endometriosis patients (derivation sample N = 1685).

VAS: Visual Analogic Scale, OCP: oral contraceptive pill, BMI: Body Mass Index, GI: gastro-intestinal, p^a:student t-test, p^b: chiz test.

select the maximum number of patients with endometriosis; (iii) and a high-risk score maximizing both the sensitivity and the specificity.¹⁹ Finally, the score was applied to the internal and the external validation samples. The performances of the scores to predict endometriosis were based on comparison of the ROC curves. The sensitivity and specificity of the different thresholds of the score were compared with those obtained from the derivation sample.

In order to avoid variations due to patient self-evaluation (pain assessment), we devised a second score (score 2) with the same method but excluding the VAS scores (except VAS for dysmenorrhea, a main symptom of the disease). The model performance of score 2 was assessed as for score 1 (ROC area, goodness of fit, and determination of three thresholds). The same statistics (sensitivity, specificity levels) were used at the same score thresholds for the internal and external validation.

All of the missing values were coded 0 by simple imputation,²⁰ as there were very few missing values (< 5%) and none correlated with the variable of interest (endometriosis). The statistical analyses were performed using STATA statistical software version 12.1 (Stata-Corp, College Station, TX, USA).

Role of funding source: There was no funding for this study. The raw data set was accessed by CC, PS, and MCLP and CC decided to submit the manuscript for publication.

Results

The initial population extracted from the database comprised 2 527 patients who were divided into two groups: (i) the study group of 1 195 histologically proven endometriosis patients; (ii) the control group of 1 332 patients who did not have any endometriotic lesions during surgery. The patient distribution according to their worst endometriotic lesions was: SPE (234; 19.6%), OMA (310; 25.9%), and DIE (651; 54.5%). Of the 651 DIE patients, 227 (34.9%) exhibited an associated OMA. The external validation cohort from Russia comprised 308 patients, including 88 controls and 220 with histologically proven endometriosis recruited as follows: SPE (88; 40.0%), OMA (88; 40.0%), and DIE (44; 20.0%).¹⁴

After a random selection of two-thirds of our population, we obtained 1 685 patients in the derivation cohort: 885 served as control patients and 800 had endometriosis. The patient characteristics are presented in [Table 1](#). The remaining one-third of the sample comprised 842 patients for the internal validation, and it included 447 controls and 395 endometriosis patients.

After selection of variables as previously described based on statistical and clinical criteria, 14 significant variables according to the univariate analysis ([Table 2](#)) were introduced in a multiple logistic regression analysis: 10 variables that were independently associated with the diagnosis of endometriosis were selected by a backward procedure. Eight of these were selected for a simplified model: the two variables excluded were: nulliparity and blood group and rhesus. The R^2 decreased from 0.24 to 0.23 and the ROC curve of the model decreased from 0.82 to 0.81. The goodness of fit had a p-value of 0.35 for ten variables and 0.87 for 8 variables. The β coefficients from the logistic regression were rounded and multiplied by ten, and the score was then calculated according to this equation as shown in [Table 3](#):

$$\text{Score } 1 = [(\text{family history of endometriosis} \times 14) + (\text{primary infertility} \times 6) + (\text{BMI} < 22 \times 7) + (\text{regular cycles} < 28 \text{ days} \times 4) + (\text{VAS dysmenorrhea} \geq 6 \times 11) + (\text{VAS dyspareunia} \geq 3 \times 6)] + (\text{VAS GI symptoms} \geq 5 \times 14) + (\text{VAS urinary tract symptoms} \geq 1 \times 12).$$

The aOR and the β coefficients from the logistic regression are presented in [Table 3](#).

The C-index of the simplified score for the diagnosis of endometriosis was 0.81 (95% CI [0.79–0.83]) ([Fig. 1a](#)) for the derivation sample. The calibration of the model was evaluated using the Hosmer-Lemeshow test with a p-value of 0.87. This score was validated on the internal validation sample [0.81 (95% CI [0.78–0.84])] and in the external validation sample [0.73 (95% CI [0.67–0.79])] ([Fig. 1b](#)).

We then defined three levels for score 1 in the derivation sample: a “low-risk group” (score 1 < 11), resulting

Variables	Se (%)	Sp (%)	C index	OR [95%CI]	p
VAS pains scores					
VAS dysmenorrhea ≥ 6	75.6 [72.5–78.6]	62.8 [59.5–66.0]	0.71	5.2 [4.2–6.5]	<0.001
VAS deep dyspareunia ≥ 3	67.5 [64.1–70.7]	62.4 [59.1–65.6]	0.67	3.4 [2.8–4.2]	<0.001
VAS GI symptoms ≥ 5	58.8 [55.2–65.2]	91.7 [93.5–95.1]	0.65	8.5 [6.3–11.6]	<0.001
VAS urinary tract symptoms ≥ 1	10.5 [8.4–12.8]	92.5 [90.6–94.2]	0.60	6.5 [4.3–9.6]	<0.001
Absence from school because of primary dysmenorrhea or loss of consciousness due to dysmenorrhea	17.5 [14.9–20.3]	95.3 [93.6–96.6]	0.58	2.5 [2.0–3.7]	<0.001
OCP for severe primary dysmenorrhea	22.5 [19.6–25.6]	91.9 [89.9–93.6]	0.59	3.3 [2.4–4.4]	<0.001
Rectorrhagia or hematuria	17.5 [14.9–20.3]	95.3 [93.6–96.6]	0.58	4.3 [3.0–6.1]	<0.001
Birth weight < 2500 g	6.25 [4.67–8.7]	96.5 [95.1–97.6]	0.52	1.8 [1.2–2.9]	0.009
Nulliparity	82.4 [79.6–84.0]	28.0 [25.1–31.1]	0.57	1.8 [1.4–2.3]	<0.001
Primary sterility	26.3 [23.2–29.4]	82.0 [79.3–84.5]	0.56	1.6 [1.3–2.0]	<0.001
BMI < 22	60.3 [56.8–63.7]	55.4 [52.0–58.7]	0.60	1.9 [1.5–2.5]	<0.001
Negative Rhesus	10.5 [8.46–12.8]	92.5 [90.6–94.2]	0.53	1.5 [1.04–2.04]	0.03
Family history of endometriosis	10.9 [8.8–13.2]	97.9 [96.7–98.7]	0.56	5.6 [3.3–9.2]	<0.001
Short regular cycles (<28 days)	9.7 [7.8–12.0]	91.4 [89.4–93.2]	0.51	1.15 [0.83–1.60]	0.20

Table 2: Diagnostic performances of dichotomized variables associated with endometriosis in the derivation sample N = 1685.

VAS: Visual Analogic Scale, OCP: oral contraceptive pill, BMI: Body Mass Index, GI: gastro-intestinal, Se: sensitivity, Spe: specificity, OR: crude odd ratio.

Variables	Score 1		Score 2		Score 1 calculation		Score 2 calculation	
	ORa [95% CI]	Simplified β coefficients	ORa [95% CI]	Simplified β coefficients	yes	no	yes	No
Family history of endometriosis	3.99 [2.3–7.1]	1.4	5.3 [3.1–9.0]	1.7	14	0	17	0
Primary sterility	1.9 [1.4–2.5]	0.6	1.7 [1.3–2.2]	0.5	6	0	5	0
BMI < 22	1.9 [1.5–2.4]	0.7	5.3 [4.3–6.6]	0.7	7	0	7	0
Short regular cycles (<28 days)	1.6 [1.1–2.4]	0.4	1.3 [1.0–1.9]	0.2	4	0	2	0
VAS pains scores								
VAS dysmenorrhea ≥ 6	3.1 [2.4–3.9]	1.1	5.3 [3.1–9.0]	1.7	11	0	17	0
VAS deep dyspareunia ≥ 3	1.92 [1.5–2.4]	0.6			6	0		
VAS GI symptoms ≥ 5	4.0 [1.9–5.6]	1.4			14	0		
VAS urinary tract symptoms ≥ 1	3.2 [2.0–5.0]	1.2			12	0		
					Sum= score1	Sum= score2		

Table 3: Association between prediction variables and endometriosis after selection: Final simplified model with 8 variables (score 1) and five variables (score 2) and calculation of the scores.
 VAS: Visual Analogic Scale, OCP: oral contraceptive pill, BMI: Body Mass Index, GI: gastro-intestinal, Se : sensitivity, Spe : specificity, AUC : area under the curve ORa : adjusted odd ratio.

in a sensitivity of 91% (95% CI [89–93]); a “high-risk group” (score 1 ≥ 18), resulting in a sensitivity of 73% (95% CI [70–76]) and a specificity of 75% (95% CI [72–78]); and a “very high-risk group” (score 1 ≥ 25), with a specificity of 91% (95% CI [89–93]) (Table 4a). All of these results were validated on the internal validation population (Table 4a). In the external validation

population, the specificity was in the expected range for the high group and the very high group but the sensitivity was slightly lower for the low-risk group (Table 4a).

We noticed that in the derivation population the high-risk level of score 1 (≥ 25) selected a sample comprising 84% (95% CI [81–87]) endometriosis patients and 16% (95% CI [13–19]) controls. The low-risk level (< 11) selected

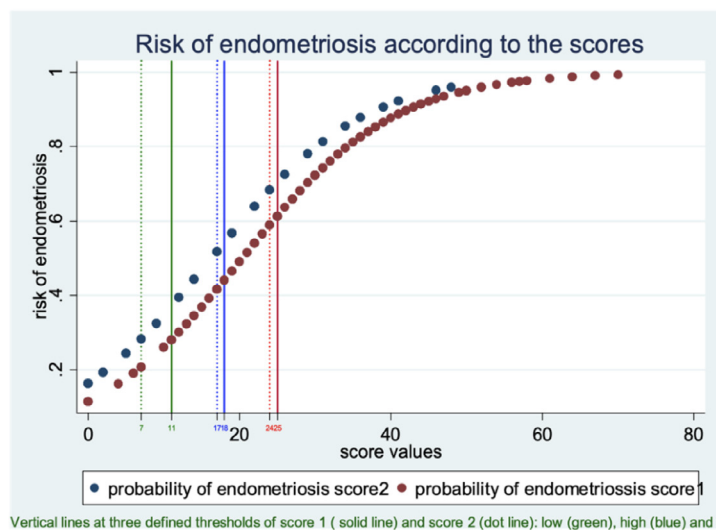


Fig. 1a. Endometriosis risk curves based on scores 1 and 2.

The curves show the relationship between the score values (x-axis) and the risk of endometriosis (y-axis). The curves for scores 1 and 2 are presented in the same figure. The theoretical risk of endometriosis is calculated from the model based on the results of regression analysis. The score was calculated for each patient with the coefficients of the multiple regression.

Above a value of 25 for score 1 and a value of 24 for score 2, the risk of endometriosis was greater than 60% and increased rapidly (most of the endometriosis patients were in this group). Below a value of 11 for score 1 and a value of 7 for score 2, the risk was less than 30% (most of disease-free patients were in this group). Between these two values, the risk increased progressively. The intermediate value of 18 for score 1 and a value of 17 for score 2 provided the best cut-off point to classify the whole population.

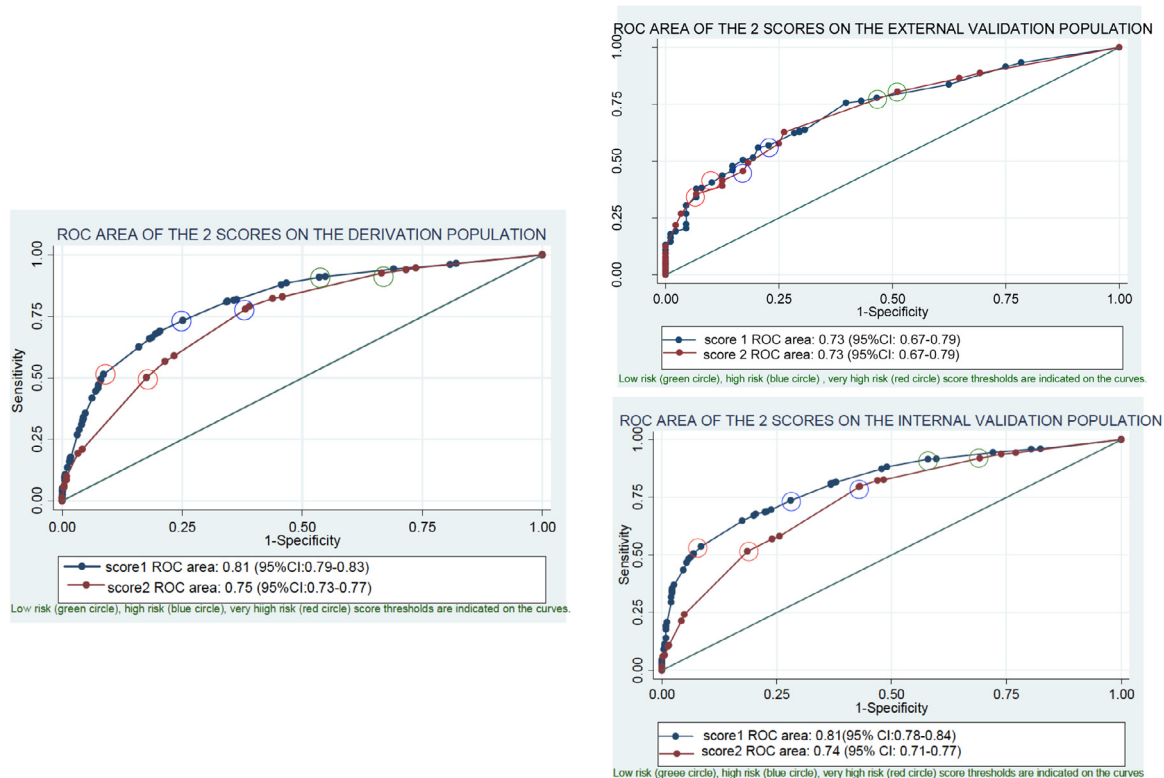


Fig. 1b. Performance of score 1 and score 2 according to the ROC area for the prediction of endometriosis on the derivation population validated on the internal and external populations.

The performance of the scores to predict endometriosis was evaluated by the ROC area.

For score 1, we used 8 variables, 4 of which were related to the patient's pelvic pain self-evaluation. The ROC area of the derivation sample exhibited good performances and was not statistically different in the internal and the external validation populations.

For score 2, we used 5 variables, with only one related to the patient's pelvic pain self-evaluation. The ROC area decreased slightly but remained good and was validated on the internal and the external validation populations.

We used a circle to highlight the score values to select a population at very high risk (high specificity: red circle), low risk (high sensitivity: green circle), and intermediate risk, which optimizes both the sensitivity and the specificity (blue circle).

a sample comprising 85% (95% CI [82–88]) controls and 15% (95% CI [12–18]) endometriosis patients.

For the whole population of endometriosis patients, the high-risk level of score 1 (≥ 25) selected a sample comprising 51% (95% CI [47–55]) endometriosis cases, including 69% (95% CI [66–70]) of the DIE cases in our population, while the low-risk level of score 1 (< 11) selected a sample comprising 9% (95% CI [6–12]) endometriosis patients, including only 3% (95% CI [1–5]) of the DIE cases.

For the entire control population, the low-risk level of score 1 (< 11) selected a sample comprising 46% (95% CI [43–49]) disease-free patients, while the high-risk level of score 1 (≥ 25) selected a sample comprising 9% (95% CI [7–11]) disease-free patients.

Score 2, with 5 variables, was obtained by the same method according to this equation, as shown in Table 3:

Score 2 = [(family history of endometriosis $\times 17$) + (primary infertility $\times 5$) + (BMI $< 22 \times 7$) + (regular cycles < 28 days $\times 2$) + (VAS dysmenorrhea $\geq 6 \times 17$).

Score 2 had a C-index of 0.75 (95% CI [0.73–0.77]) in the derivation population (Fig. 1b), validated on the internal validation (0.74 (95% CI [0.71–0.77])) and on the external validation cohorts (0.73 (95% CI [0.67–0.79])) (Fig. 1b). The low-risk level (score 2 < 7) had a sensitivity of 92% (95% CI [90–94]). The high-risk level (score 2 ≥ 17) had a sensitivity of 78% (95% CI [75–81]) and a specificity of 62% (95% CI [58–65]). The very high-risk level (≥ 24) had a specificity of 82% (95% CI [80–85]) (Table 4b). Score 2 was validated on the internal validation population: it exhibited better performances in terms of specificity in the external validation population, but its sensitivity was lower (Table 4b).

Discussion

We have developed and validated (internally and externally) the overall performances of two clinical scores that predict the probability of endometriosis (Tables 3, IVa,b, and Figs. 1a,b).

SCORE 1	Se [95%CI]	Spe [95%CI]	Observed% [95%CI] of population of endometriosis patients	Observed% [95%CI] of population of controls	Observed% [95%CI] in the sample selected by the value of score 1
DERIVATION SAMPLE N = 1675 (880 endometriosis, 802 controls)					
<11	91	46	9% [6–12]	46% [43–49]	15% endometriosis
N = 482	[89–93]	[43–50]	(3% [1–5] DIE)		85% controls
≥18	73	75	73% [70–76]	25% [22–28]	72% endometriosis
N = 806	[70–76]	[72–78]	(84% [81–87] DIE)		27% controls
≥25	51	91	51% [47–55]	9% [7–11]	84% endometriosis
N = 486	[48–55]	[89–93]	(69% [66–70] DIE)		16% controls
INTERNAL VALIDATION SAMPLE N = 842 (395 endometriosis, 447 controls)					
<11	91	42	9% [6–12]	42% [37–47]	15% endometriosis
N = 221	[88–94]	[37–45]	(1% [0–2] DIE)		85% controls
≥18	74	72	74% [70–78]	28% [24–32]	70% endometriosis
N = 415	[69–78]	[68–76]	(89% [85–93] DIE)		30% controls
≥25	54	92	54% [49–59]	9% [6–12]	85% endometriosis
N = 249	[49–59]	[89–94]	(74% [68–80] DIE)		15% controls
EXTERNAL VALIDATION SAMPLE N = 308 (220 endometriosis, 88 controls)					
<11	78*	53	22% [17–27]	53% [43–63]	NA**
N = 96	[72 83]	[42 64]	(7% [0–14] DIE)		
≥18	57*	77	57% [50–64]	23% [14–32]	NA**
N = 145	[50–64]	[67–86]	(73% [70–86] DIE)		
≥25	41*	90	40% [34–46]	10% [4–16]	NA**
N = 98	[34 47]	[82–95]	(59% [44–64] DIE)		

Table 4a: Comparison of the performances of the low risk (score 1 < 11), high risk (score 1 ≥ 18) and very high risk (score 1 ≥ 25) thresholds of the score 1 to select endometriosis patients in the three populations (derivation, internal and external validation).

CI: confidence interval, Se: sensitivity, Spe: specificity.

* statistically different from the value in the derivation sample ($p < 0.05$), NA.

** non adapted for a case control study.

Score 1, using 8 variables, predicted the occurrence of endometriosis, with an ROC area of 0.81 (95% CI [79–83]) validated on the internal (C-index of 0.81 (95% CI [78–84])) and the external (C-index of 0.73 (95% CI [67–79])) validation populations.

We chose three cut-off points: (i) score 1 ≥ 25, with a very high specificity (91%, 95% CI [89–93]); (ii) score 1 < 11, with a high sensitivity (91%, 95% CI [89–93]); (iii) and score 1 ≥ 18, maximizing both the sensitivity (73%, 95% CI [70–76]) and the specificity (75%, 95% CI [72–78]). All performances were validated on the internal validation population. The specificity of the high (score 1 ≥ 18) and the very high (score 1 ≥ 25) risk groups were validated on the external validation cohort.

Score 2 avoided parameters subject to variations due to patient self-evaluations. With 5 variables, score 2 had a C-index of 0.75 (95% CI [73–77]), validated on the internal (C-index of 0.74 (95% CI [71–77])) and the external (C-index of 0.73 (95% CI [67–79])) validation populations. The C-index was not statistically different in the external population from those of score 1, while the performances decreased in the derivation and the internal validation sample.

For both scores, we noticed that the selection was better for the most severe cases (i.e., DIE patients). These results suggest that the better selection of DIE

was due not only to the VAS scores but also to other variables.

The strength of the study lies with the large number of patients, the accurate diagnosis based on histology and surgical exploration, the prospective collection of data minimizing the number of missing data and bias, and the external validation of the score in addition to the internal validation, which has never been done before. However, our study also has some weaknesses and limitations: (i) patients who had surgery may not be representative of the entire population of patients with endometriosis, and (ii) the controls had gynecological surgical pathologies, which could mean that they were different from the general population. Unfortunately, all the other selection methods that have been used to date have their own biases and disadvantages that can interfere even more negatively in a diagnostic methodology.²¹ Patients undergoing surgery may not represent the broader population of endometriosis-affected women. However, a surgical population with histological proof allows us to be certain of the endometriotic nature of the lesions in patients and the absence of silent endometriosis in the controls. The high prevalence of GI lesions in the patients undergoing surgery for endometriosis in our department may have increased the β coefficient of the GI symptoms.

SCORE 2	Se [95%CI]	Spe [95%CI]	Observed% [95%CI] of population of endometriosis patients	Observed% [95%CI] of population of controls	Observed% [95%CI] in the sample selected by the value of score 2
DERIVATION SAMPLE N = 1682 (880 endometriosis, 802 controls)					
<7	92	33	7% [5–9]	33%	14% endometriosis
N = 355	[90–94]	[30–36]	(3% [1–5] DIE)	[30–36]	86% controls
≥17	78	62	79% [76–82]	38%	65% endometriosis
N = 721	[75–81]	[58–65]	(88% [85–91] DIE)	[35–41]	35% controls
≥24	50	82	50% [47–53]	18%	78% endometriosis
N = 558	[47–54]	[80–85]	(54% [49–59] DIE)	[16–20]	22% controls
INTERNAL VALIDATION SAMPLE N = 842 (395 endometriosis, 447 controls)					
<7	92	31	8% [5–11]	31%	19% ose
N = 169	[89–94]	[26–35]	(4% [1–7] DIE)	[27–35]	81% controls
≥17	79	57	79% [75–83]	43%	62% endometriosis
N = 507	[75–83]	[52–61]	(88% [83–93] DIE)	[38–48]	38% controls
≥24	52	81	52% [47–57]	18%	71% endometriosis
N = 287	[47–57]	[78–85]	(59% [52–66] DIE)	[14–22]	29% controls
EXTERNAL VALIDATION SAMPLE N = 308 (220 endometriosis, 88 controls)					
<7	81*	49	19% [14–23]	49%	NA**
N = 86	[75–86]	[38–60]	(14% [4–24] DIE)	[39–59]	
≥17	45*	83	45% [38–52]	17%	NA**
N = 115	[39–52]	[73–90]	(64% [50–78] DIE)	[9–25]	
≥24	35*	93	35% [29–41]	7%	NA**
N = 84	[29–42]	[86–98]	(48% [43–53] DIE)	[2–12]	

Table 4b: Comparison of the performances of the low risk (score2<11), high risk (score2≥17) and very high risk (score2≥24) thresholds of the score 2 to select endometriosis patients in the three populations (derivation, internal and external validation).

CI: confidence interval, Se: sensitivity, Spe: specificity.

* statistically different from the value in the derivation sample ($p < 0.05$).

** case control study, NA^{**}: non adapted for a case control study.

Moreover, the number of patients required to prove the absence of a difference between the derivation and the external cohort may have been too small. However, the diagnostic performance of the score applied to the external control cohort was nonetheless good and significant.

As a patient questionnaire is a key component in treating endometriosis, many questionnaires have been developed to evaluate the disease. For example, the Medical Outcomes Study Short Form 36 (SF-36),²² the Endometriosis Health Profile 30 (EHP-30),²³ and the EHP-5 questionnaire²⁴ are used to measure the impact of endometriosis on health-related quality of life (HrQoL) and/or to assess treatment efficiency, and they have been validated²⁵ and are useful in the daily practice.²⁶ The goal for other studies has been to measure the painful symptoms of endometriosis patients²⁷ or to predict rectal involvement.²⁸ Asking patients about their adolescent history can identify markers associated with deep lesions in endometriosis patients.²⁹ A clinical score based solely on a patient clinical history can predict DIE before surgery for patients with OMA diagnosed by TVUS,¹³ and a standardized evaluation of painful symptoms allows a satisfactory rate of detection of women with posterior DIE.³⁰

Nevertheless, despite the difficulty diagnosing endometriosis, very few logistic regression models have been devised to predict endometriosis. Stegmann et al.³¹ developed a model that correlated the individual lesion characteristics with the pathological findings in a multivariate analysis. This model has a modest capacity to predict endometriosis, and the main limitation is that histological diagnoses are used in the logistic regression, which necessitates performing a laparoscopy. More recently, Nnoaham et al.³² performed a prospective, observational, two-phase multicenter study with the objective of validating symptoms-based models (using a 25-item questionnaire) to predict endometriosis in symptomatic women who had not previously undergone a laparoscopy. The ability to predict any endometriosis stage in the model was relatively poor (area under the curve = 0.683). The predictive value of the model was improved by ultrasound scan evidence of cysts/nodules. To the best of our knowledge, our score is the first one based solely on patient clinical history that provides a satisfactory statistical performance for predicting endometriosis with an external validation.

We are at a key stage regarding the diagnosis and management of endometriosis, both of which warrant being extensively reconsidered.¹ As histological

diagnosis is no longer required to initiate treatment (s),⁸⁻¹² there is a high level of interest in non-invasive (i.e., non-surgical) ways to diagnose endometriosis.³ The selection of patients requiring a radiological evaluation based on a simple patient questionnaire is crucial for the medical care. Very high-risk patients need to be referred to a multidisciplinary specialized team (expert radiologists, multidisciplinary surgical team (gynecological, GI, urological, and thoracic surgeons), assisted reproductive technology (ART) specialists, gynecological endocrinologists, pathologists, physiotherapists, psychologists, etc.). If a patient is at high risk with a negative radiological examination, SPE cannot be excluded, and the patient is managed accordingly for suspected endometriosis. This new paradigm regarding endometriosis diagnosis, based on the patient's clinical history and imaging, allows use of a non-surgical procedure to devise an individualized therapeutic strategy (medical treatment, surgery, or ART) for each patient. In selected cases, medical treatment and ART can be satisfactory alternatives to surgery, which is usually proposed too often as the first therapeutic option.¹

In conclusion, we have developed two internally and externally validated scores based solely on 8 and 5 items, respectively, from a patient questionnaire that reliably predict the probability of endometriosis. These scores may have high clinical value, allowing the selection of patients at high-risk of endometriosis with a cost-effective procedure that can readily be performed by the patient herself. These scores will contribute to shortening the delay of the diagnosis and improvement of the management, wellbeing, and quality of life of patients.³³ The accumulation of data is likely to help progressively improve the performance of this straightforward and cost-effective screening score.

Data sharing

Data supporting the study are available from the correspondent author (CC) upon reasonable request.

Contributors

CC coordinated the project. CC and MCLP designed the study and supervised the drafting and redrafting of the manuscript. All of the authors contributed to the data collection. CC, MCLP, BB, PS, AGC, and LM performed the surgical procedures. MCLP, who is an expert in statistics, supervised and reviewed the statistical analysis. All of the authors contributed substantially to the analysis and the interpretation of the data. The raw data set was accessed by CC, PS, and MCLP. All of the authors contributed to writing the manuscript. CC, PS, and MCLP critically revised the final version of the manuscript. All of the authors approved the final submitted version.

Declaration of interests

None to declare.

Funding

None.

Acknowledgments

The authors thank the staff members of our surgical unit for their expert assistance with the data collection. The authors also acknowledge Jeanne Colombe-Debat for continuously managing the patient database. The authors thank Ipsen for providing anonymized clinical data from the FEELING study (NCT01351051),¹⁴ which was used as the external validation cohort. Ipsen provided a courtesy review of the manuscript but had no involvement in this work.

References

- 1 Chapron C, Marcellin L, Borghese B, Santulli P. Rethinking mechanisms, diagnosis and management of endometriosis. *Nat Rev Endocrinol* 2019.
- 2 Taylor HS, Kotlyar AM, Flores VA. Endometriosis is a chronic systemic disease: clinical challenges and novel innovations. *Lancet* 2021;397:839-52.
- 3 Agarwal SK, Chapron C, Giudice LC, et al. Clinical diagnosis of endometriosis: a call to action. *Am J Obstet Gynecol* 2019;220:354.e1-e12.
- 4 Nnoaham KE, Hummelshoj L, Webster P, et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertil Steril* 2011;96:366-73. e8.
- 5 Pluchino N, Wenger JM, Petignat P, et al. Sexual function in endometriosis patients and their partners: effect of the disease and consequences of treatment. *Hum Reprod Update* 2016;22:762-74.
- 6 Guerriero S, Alcazar JL, Pascual MA, et al. Deep infiltrating endometriosis: comparison between 2-dimensional ultrasonography (US), 3-dimensional US, and magnetic resonance imaging. *J Ultrasound Med* 2018;37:1511-21.
- 7 Van den Bosch T, de Bruijn AM, de Leeuw RA, et al. A sonographic classification and reporting system for diagnosing adenomyosis. *Ultrasound Obstet Gynecol* 2018. <https://doi.org/10.1002/uog.19096>.
- 8 Dunselman GA, Vermeulen N, Becker C, et al. ESHRE guideline: management of women with endometriosis. *Hum Reprod* 2014;29:400-12.
- 9 Leyland N, Casper R, Laberge P, Singh SS. Sogc. Endometriosis: diagnosis and management. *J Obstet Gynaecol Canada: JOGC = Journal d'obstetrique et gynecologie du Canada: JOGC* 2010;32:S1-32.
- 10 American College of Obstetricians and Gynecologists. ACOG: practice bulletin no. 114: management of endometriosis. *Obstet Gynecol* 2010;116:223-36.
- 11 Johnson NP, Hummelshoj L. World endometriosis society montpellier C. Consensus on current management of endometriosis. *Hum Reprod* 2013;28:1552-68.
- 12 National Institute for Health and Care Excellence (NICE). Diagnosis and management of endometriosis: summary of NICE guidance. *BMJ* 2017;358:4227.
- 13 Lafay Pillet MC, Huchon C, Santulli P, Borghese B, Chapron C, Fauconnier A. A clinical score can predict associated deep infiltrating endometriosis before surgery for an endometrioma. *Hum Reprod* 2014;29:1666-76.
- 14 Chapron C, Lang JH, Leng JH, et al. Factors and regional differences associated with endometriosis: a multi-country, case-control study. *Adv Ther* 2016;33:1385-407.
- 15 Dawood MY. Primary dysmenorrhea: advances in pathogenesis and management. *Obstet Gynecol* 2006;108:428-41.

- 16 Efron B, Gong G. A leisurely look at the bootstrap, the Jackknife, and cross-validation. *Am Stat* 1983;37:36–48.
- 17 Lemeshow S, Hosmer Jr. DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;115:92–106.
- 18 Coste J, Bouyer J, Job-Spira N. Construction of composite scales for risk assessment in epidemiology: an application to ectopic pregnancy. *Am J Epidemiol* 1997;145:278–89.
- 19 Buckley RG, King KJ, Disney JD, Ambroz PK, Gorman JD, Klausen JH. Derivation of a clinical prediction model for the emergency department diagnosis of ectopic pregnancy. *Acad Emergency Med* 1998;5:951–60.
- 20 Pedersen AB, Mikkelsen EM, Cronin-Fenton D, et al. Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol* 2017;9:157–66.
- 21 Zondervan KT, Cardon LR, Kennedy SH. What makes a good case-control study? Design issues for complex traits such as endometriosis. *Hum Reprod* 2002;17:1415–23.
- 22 Jones GL, Kennedy SH, Jenkinson C. Health-related quality of life measurement in women with common benign gynecologic conditions: a systematic review. *Am J Obstet Gynecol* 2002;187:501–11.
- 23 Jones G, Kennedy S, Barnard A, Wong J, Jenkinson C. Development of an endometriosis quality-of-life instrument: the endometriosis health profile-30. *Obstet Gynecol* 2001;98:258–64.
- 24 Jones G, Jenkinson C, Kennedy S. Development of the Short Form Endometriosis Health Profile Questionnaire: the EHP-5. *Qual Life Res* 2004;13:695–704.
- 25 Stull DE, Wasiak R, Kreif N, et al. Validation of the SF-36 in patients with endometriosis. *Qual Life Res* 2014;23:103–17.
- 26 Comptour A, Pereira B, Lambert C, et al. Identification of predictive factors in endometriosis for improvement in patient quality of life. *J Minim Invasive Gynecol* 2019.
- 27 Fauconnier A, Staraci S, Darai E, et al. A self-administered questionnaire to measure the painful symptoms of endometriosis: results of a modified DELPHI survey of patients and physicians. *J Gynecol Obstetrics Human Reproduct* 2018;47:69–79.
- 28 Chattot C, Huchon C, Paternostre A, Du Cheyron J, Chouillard E, Fauconnier A. ENDIRECT: a preoperative score to accurately predict rectosigmoid involvement in patients with endometriosis. *Human Reproduction Open* 2019;2019:hoz007.
- 29 Chapron C, Lafay-Pillet MC, Monceau E, et al. Questioning patients about their adolescent history can identify markers associated with deep infiltrating endometriosis. *Fertil Steril* 2011;95:877–81.
- 30 Chapron C, Barakat H, Fritel X, Dubuisson JB, Breart G, Fauconnier A. Presurgical diagnosis of posterior deep infiltrating endometriosis based on a standardized questionnaire. *Hum Reprod* 2005;20:507–13.
- 31 Stegmann BJ, Funk MJ, Sinaii N, et al. A logistic model for the prediction of endometriosis. *Fertil Steril* 2009;91:51–5.
- 32 Nnoaham KE, Hummelshoj L, Kennedy SH, Jenkinson C, Zondervan KT. World Endometriosis Research Foundation Women's Health Symptom Survey C. Developing symptom-based predictive models of endometriosis as a clinical screening tool: results from a multicenter study. *Fertil Steril* 2012;98:692–701. e5.
- 33 Rush G, Misajon R, Hunter JA, Gardner J, O'Brien KS. The relationship between endometriosis-related pelvic pain and symptom frequency, and subjective wellbeing. *Health Qual Life Outcomes* 2019;17:123.