Human Reproduction, pp. 1–12, 2017
doi:10.1093/humrep/dex057

human reproduction

ORIGINAL ARTICLE ***Psychology & counseling***

Measuring health-related quality of life in women with endometriosis: comparing the clinimetric properties of the Endometriosis Health Profile-5 (EHP-5) and the EuroQol-5D (EQ-5D)

G. Aubry[1], P. Panel[2], G. Thiollier[3], C. Huchon[1,3], and A. Fauconnier[1,3,*]

[1]EA 7285 Research Unit 'Risk and Safety in Clinical Medicine for Women and Perinatal Health', Versailles-Saint-Quentin University (UVSQ), 78180 Montigny-le-Bretonneux, France [2]Department of Gynecology and Obstetrics, Centre Hospitalier Intercommunal de Versailles, 78150 Le Chesnay, France [3]Department of Gynecology and Obstetrics, Centre Hospitalier Intercommunal de Poissy-Saint-Germain-en-Laye, 78300 Poissy, France

*Correspondence address. Department of Gynecology and Obstetrics, Intercommunal Hospital Centre of Poissy-Saint-Germain-en-Laye, Poissy Cedex, France. Tel: (+33)-1-39-27-51-51; Fax: (+33)-1-39-27-44-12; E-mail: afauconnier@chi-poissy-st-germain.fr

Submitted on November 4, 2016; resubmitted on February 13, 2017; accepted on March 9, 2017

STUDY QUESTION: Which of the Endometriosis Health Profile-5 (EHP-5) and the EuroQol-5D (EQ-5D) is the most efficient to assess quality of life in women suffering from endometriosis?

SUMMARY ANSWER: Although EHP-5 and EQ-5D instruments had an excellent responsiveness, EHP-5 has a better discriminative ability than EQ-5 to measure health-related quality of life (HrQoL).

WHAT IS KNOWN ALREADY: Proper measurement of HrQoL is important in endometriosis. While many quality of life instruments are available, few have been completely validated in endometriosis. The EHP-5 and the EQ-5D are short and practical scales, which may be useful. Literature is lacking to determine which one is the most suitable in clinical practice or in clinical research.

STUDY DESIGN, SIZE, DURATION: This prospective and observational study conducted between 1 January 2012 and 31 December 2013 included a total of 253 consecutive women with proven endometriosis, undergoing medical or surgical treatment, in 2 French tertiary care centers.

PARTICIPANTS/MATERIALS, SETTINGS, METHODS: Women over 18 years consulting for painful symptoms of at least 3 months' duration or for infertility, with endometriosis proven histologically or radiologically, were requested to fill in the 2 scales before (T_0) and 12 months after treatment (T_1). Construct validity consisted in testing presupposed relationships between the scales and the characteristics of the patients or the endometriosis. Responsiveness to change was calculated for all patients and in each treatment group. Effect sizes were used according to Cohen's *d* method.

MAIN RESULTS AND THE ROLE OF CHANCE: A total of 216 women filled in completely all the questionnaires at T_0 and 133 (61.6%) at T_1. EHP-5 and EQ-5D had good discriminative abilities regarding the patients' symptoms, with significant superiority of EHP-5 concerning three of the nine hypotheses. The largest difference was that calculated for the 'intensity of dysmenorrhea' using the Visual Analogic Scale, with respectively effect size from Cohen's *d* (ES) = 0.86 95% CI (0.54–1.17) for EHP-5 versus 0.48 95% CI (0.16–0.79) for EQ-5D. There were no differences in EHP-5 or in EQ-5D scores between subgroups according to the characteristics of endometriosis. Overall responsiveness was excellent and equivalent for EHP-5 and for EQ-5D, with, respectively, ES = 0.81 95% CI (0.56–1.56) versus ES = 0.95 95% CI (0.68–1.20). In subgroup analyses, EHP-5 was responsive in case of medical treatment with ES = 0.93 95% CI (0.07–1.70), whereas EQ-5D was not, ES = 0.73 95% CI (−0.06–1.47).

LIMITATIONS, REASONS FOR CAUTION: Our study population included patients with symptomatic and mainly severe forms of endometriosis, which may suggest a spectrum bias. The evaluation of responsiveness in case of medical treatment was based on a small number of patients, which limits the interpretation of the difference found between the two scales in this subgroup.

WIDER IMPLICATIONS OF THE FINDINGS: EHP-5 is a simple, efficient and valid tool for evaluating quality of life in daily practice and also valuable to provide a primary outcome in clinical studies evaluating treatment efficacy.

STUDY FUNDING/COMPETING INTEREST(S): This work was funded by the Direction à la Recherche Clinique et à l'Innovation of Versailles, France. The authors have no conflicts of interest.

TRIAL REGISTER NUMBER: None.

Key words: endometriosis / quality of life / responsiveness / EQ-5D/EHP-5

Introduction

Endometriosis is the most common benign gynecological disease, the prevalence of which reaches 35–50% in women with pain, infertility or both (Giudice, 2010; Schliep *et al.*, 2015). The associated symptoms, dominated by severe pelvic pains and infertility, the chronicity, side effects of treatment and lack of understanding by other people, can have an impact on the personal, psychological and social aspects of patients' daily life (Fourquet *et al.*, 2011; De Graaff *et al.*, 2013).

Recent studies have demonstrated a lower health-related quality of life (HrQoL) in women with endometriosis compared to the general population (Nnoaham *et al.*, 2011; Simoens *et al.*, 2012). However, a systematic review conducted in 2002 emphasized a lack of HrQoL evaluation in those patients, mostly due to a lack of efficient and valid tools (Jones *et al.*, 2002). In France, endometriosis and the HrQoL of these patients are becoming one of the main concerns in gynecological health intervention policies. Thus, there is a need for an optimal and valid instrument that could be used in daily practice or in research.

The 'Endometriosis Health Profile 30' (EHP-30) is a patient-generated instrument, considered as a 'Patient-Reported Outcome' (PRO) instrument (Jones *et al.*, 2001; Turk *et al.*, 2006). It is the only existing HrQoL questionnaire specific to endometriosis. A shorter version, more practical and suitable for clinical practice and also for research, was developed: the Endometriosis Health Profile-5 (EHP-5) (Jones *et al.*, 2004a). EHP-5 has been translated into French (Renouvel *et al.*, 2009) and exhibited excellent psychometric validity (Fauconnier *et al.*, 2017). Checking on the clinimetric properties, i.e. construct validity and responsiveness, is the last step to complete the full validation process of the French version of EHP-5. Moreover, its responsiveness has never been explored, even in its original version, even though evidence supporting responsiveness is critical for daily practice as well as for clinical trial settings (Dworkin *et al.*, 2008).

Comparative evaluation of instrument performance has been reported to provide the most useful evidence for informing instrument selection (Garratt *et al.*, 2002). The EuroQol-5D (EQ-5D) which is a valid, short and reliable generic HrQoL instrument (EuroQol Group, 1990), translated into French, was the most suitable for comparison with EHP-5. It is one of the tools used most in HrQoL evaluation in endometriosis (Gao *et al.*, 2006; Simoens *et al.*, 2012), but has not been validated specifically for endometriosis.

The aim of the present study was to compare the clinical value of the EHP-5 and the EQ-5D, in order to determine which one was the more efficient to assess HrQoL of French symptomatic patients with endometriosis, in a short and simple manner.

Materials and Methods

Study design

This was a prospective and observational study conducted between the 1 January 2012 and 31 December 2013 at the Centre Hospitalier Intercommunal de Versailles (CHV) and the Centre Hospitalier Intercommunal de Poissy-Saint-Germain (CHIPS). These hospitals are French referral centers for treatment of endometriosis, with teams with extensive experience in surgery for deeply infiltrating endometriosis, including intestinal and bladder resection.

Study population and follow-up

The study population consisted of consecutive patients over 18 years old with proven endometriosis, consulting and treated for painful symptoms of at least 3 months' duration or for infertility. For women with surgical treatment, endometriosis was considered proven in case of typical macroscopic black-bluish nodule found during laparoscopy and/or in case of histological proof. For women with medical treatment and without laparoscopic diagnosis, we used strong criteria: (i) when a typical macroscopic black-bluish nodule was observed at posterior vaginal fornix examination or at cystoscopy or at rectal endoscopy (Vercellini *et al.*, 1996; Bazot *et al.*, 2009) and (ii) in case of typical aspects on magnetic resonance images, i.e. endometriomas or deep infiltrating endometriosis (DIE), described as hyperintense foci on T_1-weighted images, small hyperintense cavities on T_2-weighted images or areas corresponding to fibrosis on T_1- and T_2-weighted images (Bazot *et al.*, 2007). The exclusion criteria were patients who could not read French; severe disabilities or chronic pain; associated pelvic pathology; patient consulting for advice after treatment elsewhere; patients without surgery and having inconclusive physical or endoscopy or magnetic resonance images were also excluded.

At the first visit before treatment (T_0), participants were asked to complete a structured questionnaire on the characteristics of their symptoms, the French version of EHP-5 (Renouvel *et al.*, 2009) and the French version of EQ-5D (Chevalier and de Pouvourville, 2013). The decision to perform surgery or to choose medical therapy was based on shared medical decision-making after extensive discussion between the patient and the physician (Dunselman *et al.*, 2014). Both centers used the same guidelines for treatment. Surgery was indicated in case of medical treatment failure or in women refusing hormonal therapy, in case of diagnosis uncertainty, or in case of infertility without assisted reproductive technology indications. In other cases, medical hormonal therapy was the first-line choice. Surgical treatment

included excision or fulguration of superficial implants on the peritoneum, excision of endometriomas, resection of rectovaginal nodules with or without partial colpectomy or rectal shaving (Canis, 2007). Bowel resection and/or partial cystectomy were performed when indicated, with the help of a specialized surgeon if required. GnRH agonists were administered before and/or after surgery according to the surgeon's decision, with or without add-back therapy. Medical hormonal treatment consisted in the suppression of cyclic ovarian hormone production and elimination of menses, by combined oral or progestogen-only contraceptive pills (Dunselman *et al.*, 2014)

Twelve months after treatment (T_1), the three questionnaires were distributed again by post. The Clinical Global Impression-Improvement (CGI-I) scale was included in order to evaluate the evolution of the disease after treatment in a qualitative way, from the patient's point of view (Gerlinger *et al.*, 2010). The questionnaires were completed in a self-assessed way using paper and pencil.

The questionnaires

The EHP-5 is built in two parts, with questions referring to the 4 previous weeks: a 5-item core questionnaire about pain, control and powerlessness, emotions, social support, self-image and a 6-item modular questionnaire about work life, relation with children, sexual intercourse, medical profession, treatment and infertility. The response system consists of five levels ranged in order of severity: 'never', 'rarely', 'sometimes', 'often' and 'always' (Jones *et al.*, 2004a; Renouvel *et al.*, 2009).

The EQ-5D also consists of two parts: the first is the 'EQ-5D descriptive system' with five questions about several dimensions of HrQoL (mobility, self-care, daily activities, pain/discomfort and anxiety/depression). Each dimension can be rated at three levels: 'no problems', 'some problems' and 'major problems'. The responses to the five dimensions together lead to an aggregated index, the 'EQ-5D index', which corresponds to a health state. The second part of the scale is the 'EQ Visual Analogic Scale' (EQ-VAS) which is the respondent's own assessment of her overall health status on a thermometer-like visual scale (EuroQol Group, 1990; Chevalier and de Pouvourville, 2013).

The CGI-I is a one-question single index 'how your symptoms (pain and others) are now compared with how they were before treatment?' rated by seven answers: (i) 'Much better', (ii) 'Better', (iii) 'Somewhat better', (iv) 'No change', (v) 'Somewhat worse', (vi) 'Worse' and (vii) 'Much worse' (Gerlinger *et al.*, 2010).

Endometriosis classification

A standardized description sheet of anatomical endometriosis lesions based on the surgical and radiological data was used in both centers. It reported the locations of the endometriosis implants and the subtype of endometriosis (superficial endometriosis only, endometriomas and DIE). In case of DIE, it reported the depth of infiltration in order to define the type of endometriosis according to a DIE classification system previously published (Chapron *et al.*, 2003). In case of surgical management, the extent of the disease was also assessed, according to the standards set by the American Society of Reproductive Medicine classification (ASRM, 1997).

Statistical analysis

Our primary endpoint was the comparison of the overall responsiveness of the two scales, assuming that most patients would be improved, regardless of the type of treatment they would receive. The calculation of the number of participants required for responsiveness measurement was based on the effect size, according to Cohen's *d* method. The responsiveness of EHP-5 has never been explored but an effect size higher than 0.8 was required to make it useful. Based on the 95% CI of the effect size calculated for EHP-30 (Jones *et al.*, 2004b), 120 patients could lead to an effect size of 0.8 or more. With an expected response rate at T_1 of 60%, this led to the inclusion of at least 200 patients.

For EHP-5, the 11-item responses were summed and transformed according to the EHP-5 manual on a scale from 100 (worst possible HrQoL) to 0 (best possible HrQoL) (Jones *et al.*, 2004a). In case of missing data for one or more item, these items were omitted from the calculation; with <6 answers, the score was not computed. For the EQ-5D descriptive system, the five dimensions together represent a health state that corresponds to a utility value. The EQ-5D score (EQ-5D index) was obtained thanks to a valid algorithm using the utility weights, adapted to the French population with a time trade-off technique (Chevalier and de Pouvourville, 2013). If an item of data was missing, the score could not be computed. The 'EQ-5D index' ranged from −0.59 (state worse than death) to 1.00 (best possible health state). The 'EQ-VAS' ranged from 0 (worst imaginable health state) to 100 (best imaginable health state).

The evaluation of construct validity, which was our second outcome, was performed for all patients who answered the questionnaires at T_0. It was assessed using known group comparison, i.e. testing preestablished relationships between EHP-5 or EQ-5D and the characteristics of patients or the endometriosis. The hypothesized relationships were the following: impairment of the scale score with the intensity or the frequency of the main painful symptoms, with duration of pain, with number of painkillers used, with infertility, with prior surgery of endometriosis, with extent of endometriosis according to the surgical classification (ASRM, 1997) or with the extent of the DIE (Chapron *et al.*, 2003). For each symptom's variable, the cut-off point to define the two subgroups was the median of the responses to the corresponding question of the symptoms questionnaire.

Responsiveness to change was evaluated in patients who answered the questionnaires both at T_0 and T_1. We performed subgroup analyses in different preestablished situations: for all patients, for patients who reported any improvement according to CGI-I, for patients according to their treatment group.

To explore the magnitude of differences in scores between the groups of a given variable, we calculated the effect size according to Cohen's *d* method, for each scale (Kazis *et al.*, 1989). A positive effect size indicated that the variable deteriorated the scale score. Effect sizes around 0.2 are considered as small effects, around 0.5 as moderate and around 0.8 or more as large effects. For a given variable, if the CI of the effect size did not include 0, then the variable was considered to affect the scale significantly at $P < 0.05$. Comparisons between the three scales (EHP-5 and EQ-5D index and EQ-VAS) were based on the 95% CI of the effect size: for a given variable, if the mean effect size of a scale was not included within the 95% CI calculated for the other scale, then it differed significantly with $P < 0.05$.

The minimal clinically important difference (MCID) is the smallest difference in scores of an instrument's measure which is perceived by patients as beneficial or harmful, and which would lead the clinician to consider a change in treatment (Jaeschke *et al.*, 1989). It was determined for all the scales with the anchor-based method (Crosby *et al.*, 2003), using the CGI-I as an anchor (Gerlinger *et al.*, 2010). First, we checked the correlation between the score change of each scale and the response to the CGI-I, using Pearson's correlation coefficient (Revicki *et al.*, 2008). The MCID was then computed as the mean score change from T_0 to T_1 in the subgroup of patients who answered 'somewhat better' or 'better' on the CGI-I (Guyatt *et al.*, 1993).

All data were collected in a computerized database and analyzed by SPSS 22.0 software (SPSS Inc., Chicago, IL, USA).

Ethics

Women were not required to sign an informed consent form as there was no intervention. The 11th ethics research committee of Ile-de-France

considered that our study involved no intervention and was thus exempt from the French statute on biomedical research (modified version of Law 2004–806, dated 9 August 2004). We complied with all French statutes concerning patient data, confidentiality and restrictions. The study was approved by the French National Committee for Information Technology and Individual Liberties (No. 906 253).

Results

The process of our data collection is shown in Fig. 1. Two hundred and sixteen patients met the inclusion criteria and completed the T_0 questionnaires. Table I presents the demographic and disease characteristics of the participating patients: 133/216 (61.6%) completed the T_1 questionnaires after treatment. Comparison of the characteristics of the respondents versus the non-respondents showed no significant differences (mean age: 33.4 ± 6.6 versus 32.9 ± 6.8 years, Student's *t*-test, $P = 0.57$; infertility $N = 50$ versus $N = 37$, Chi^2-test, $P = 0.25$; prior surgery for endometriosis $N = 50$ versus $N = 39$, Chi^2-test, $P = 0.19$; ASRM Classification Stages I–II, $N = 39$ versus $N = 16$, Chi^2-test, $P = 0.22$). Of note, 130/216 (60.2%) answered the CGI-I: 1 (0.8%) reported 'much worse', 3 (2.3%) 'worse', 1 (0.8%) 'somewhat worse', 6 (4.6%) 'no change', 11 (8.5%) 'somewhat better', 27 (20.8%) 'better' and 81 (62.3%) 'much better'.

The construct validity of the scales is given in Table II. EHP-5 and EQ-5D had good discriminative ability according to the majority of the pain descriptors. Compared to the EQ-5D descriptive system, EHP-5

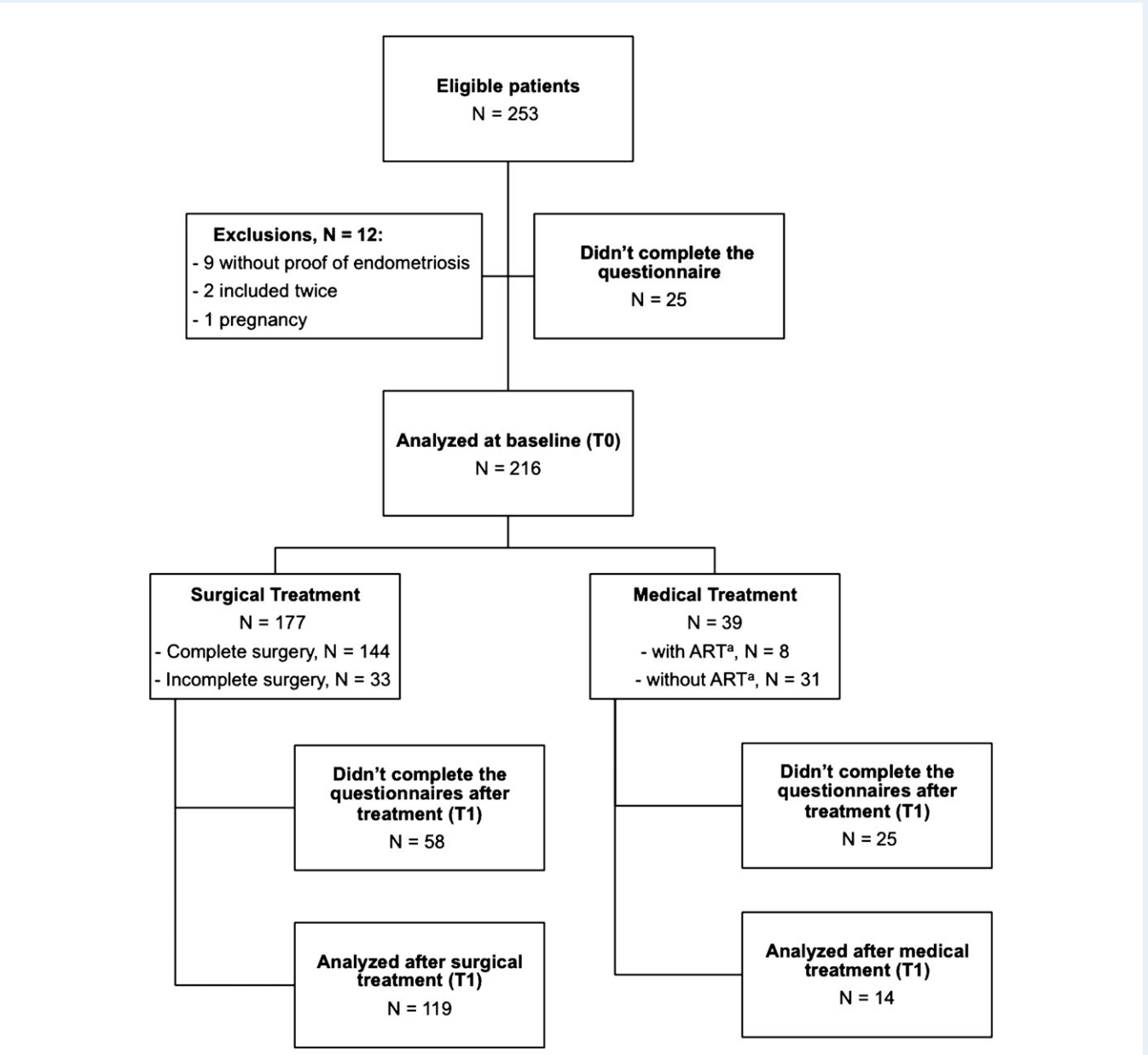

Figure 1 Flow chart of the study participants. ART, assisted reproductive technique.

Table I Demographic features and disease characteristics of the study participants ($N = 216$).

Characteristic	N (%)	Mean ± SD
Center		
Center 1 (CHV)	128 (59.3)	
Center 2 (CHIPS)	88 (40.7)	
Age (range 18–52 years)		33.2 ± 6.7
BMI (range 16.3–42.2 kg/m^2)		23.2 ± 4.7
Gravidity		
Nulligravida	124 (57.4)	
G ≥ 1	92 (42.6)	
Parity		
Nulliparous	148 (68.5)	
P ≥ 1	68 (31.5)	
Prior pelvic surgery	136 (63.0)	
For endometriosis	89 (65.4)	
Others	47 (34.6)	
Infertility	85 (39.4)	
Proof of endometriosis		
Histologic	169 (78.2)	
Macroscopic	11 (5.1)	
MRI	36 (16.7)	
ASRM surgical stage ($N = 162$)[a]		
I – Minimal	33 (20.4)	
II – Mild	23 (14.2)	
III – Moderate	52 (32.1)	
IV – Severe	54 (33.3)	
Type of endometriosis[b]		
Anterior DIE ($N = 213$)		
None	166 (77.9)	
Bladder	47 (22.1)	
Posterior DIE ($N = 209$)		
None	20 (9.6)	
Uterosacral ligament	95 (45.4)	
Vagina	11 (5.3)	
Intestine	83 (39.7)	
Endometrioma	80 (37.0)	
Main pain ($N = 205$)		
VAS ≤ 7	128 (62.4)	
VAS > 7	77 (37.5)	

VAS, Visual Analogic Scale; DIE, deep infiltrating endometriosis; ASRM, American Society of Reproductive Medicine (ASRM, 1997); CHV, Centre Hospitalier Intercommunal de Versailles; CHIPS, Centre Hospitalier Intercommunal de Poissy-Saint-Germain.

[a]Reported only for the patients with surgical treatment.

[b]Reported for all patients, according to the anatomical classification of DIE (Chapron *et al.*, 2003).

exhibited a significantly better ability to discriminate patients for the following variables: intensity of dysmenorrhea measured by a VAS, frequency of dyspareunia and number of painkillers used. For EQ-5D, patients with infertility had significantly better scores than those without infertility, as shown by the small but significant negative effect size. For EHP-5 there was no significant difference between these subgroups. Finally, there were no differences in EHP-5 or in EQ-5D scores between subgroups according to ASRM stage or DIE classification.

Table III shows the responsiveness of the scales. For the overall population of patients who responded to the T_1 questionnaires, the responsiveness of both EHP-5 and EQ-5D was similar with large effect sizes. According to treatment group, EHP-5 was the only scale to be responsive in case of medical treatment, whereas for surgical treatment, the responsiveness of EHP-5 and EQ-5D was excellent, without statistically significant difference. Table III also shows that EHP-5 and EQ-5D were even more responsive when patients reported they were better according to the CGI-I but without statistical differences between the two scales, except in case of medical treatment where EHP-5 was still the only scale to be responsive.

Change scores between T_0 and T_1 were significantly and moderately correlated with the women's response to CGI-I for EHP-5 ($r = 0.28$, $P = 0.002$) and for EQ-VAS ($r = 0.37$, $P < 0.001$) but no significant correlation was found for the EQ-5D descriptive system ($r = 0.07$, $P = 0.431$). Box plots in Fig. 2 show the variation of the scale score change between T_0 and T_1 for the 37 patients who reported they were 'somewhat better' or 'better' after treatment. The mean score change corresponds to the MCID. For EHP-5, MCID was −4.5 (standard error = 2.8); for EQ-VAS and EQ-5D index, respectively, MCID were 10.2 (standard error = 4.1) and 0.26 (standard error = 0.05).

Discussion

We found that both EHP-5 and EQ-5D exhibited excellent and equivalent responsiveness to change. However, EHP-5 was more sensitive than EQ-5D to the presence of pain symptoms that are known to impair the HrQoL of women affected by endometriosis.

The strength of this study lies in its methodological design, which was specifically dedicated to evaluate the clinimetric properties of the different HrQol questionnaires for endometriosis patients, and was not derived from secondary analysis data of previously published research clinical trials. Next, we hypothesized that most patients would be improved regardless of the treatment they had received. This made possible the study of responsiveness in a large cohort of patients with endometriosis and in a lesser extent, the evaluation of response properties among patients who benefited either from surgery or from medical therapy.

There are several limitations in our study. First, there is a possible spectrum bias since the patients included in the study had the most severe forms of endometriosis, including patients with rectal endometriosis that are the ones with the greatest impairment of HrQol (Dubernard *et al.*, 2006). This may explain the fact that we were not able to relate the HrQol and any of the variables measuring the extent of the disease, unlike previous studies including ours (Fauconnier *et al.*, 2017). Second, the number of patients in the medical treatment group is dramatically low since only 39 women chose this option, whereas medical hormonal therapy could be a first-line choice. One explanation is that we chose to use very reliable diagnostic criteria in order to minimize the number of false positives, decreasing substantively the number of patients with medical treatment eligible for the study. Another explanation is that the study took place in two referral

Table II **Construct validity of Endometriosis Health Profile-5 (EHP-5) and the EQ-5D: differences between scores of known groups predefined according to the type of symptoms or to the type of endometriosis.**

	EHP-5					EQ-5D descriptive system					EQ-VAS				
Known groups	***N*[a]**	**Mean ± SD**	***P*-value**	**Effect size**	**95% CI effect size**	***N*[a]**	**Mean ± SD**	***P*-value**	**Effect size**	**95% CI effect size**	***N*[a]**	**Mean ± SD**	***P*-value**	**Effect size**	**95% CI effect size**
Main pain intensity[b]															
≤7	127	43.5 ± 19.6	<0.001	0.62	0.33–0.91	122	0.67 ± 0.25	<0.001	0.63	0.33–0.92	119	61.5 ± 19.0	<0.001	0.69	0.38–0.99
>7	77	55.7 ± 19.9				75	0.50 ± 0.30				68	48.5 ± 18.7			
Intensity of dysmenorrhea[b]															
≤7	93	39.7 ± 20.9	<0.001	0.86	0.54–1.17	89	0.67 ± 0.25	0.003	0.48	0.16–0.79	91	63.8 ± 18.5	<0.001	0.73	0.40–1.06
>7	76	56.4 ± 17.5				72	0.54 ± 0.30				64	50.6 ± 17.4			
Intensity of pelvic pain[b]															
≤5	119	42.6 ± 21.2	<0.001	0.65	0.37–0.93	112	0.66 ± 0.27	0.001	0.47	0.19–0.74	109	61.3 ± 19.3	0.001	0.51	0.22–0.79
>5	94	55.5 ± 17.9				93	0.53 ± 0.29				85	51.5 ± 19.5			
Frequency of painful defecation															
None-mild	107	43.4 ± 21.2	<0.001	0.60	0.33–0.87	101	0.66 ± 0.30	0.004	0.40	0.12–0.67	99	62.2 ± 18.5	<0.001	0.55	0.26–0.83
Moderate-severe	108	53.4 ± 10.2				106	0.54 ± 0.30				97	51.6 ± 19.9			
Frequency of dyspareunia															
Never-sometimes	108	42.1 ± 21.0	<0.001	0.69	0.39–0.98	103	0.64 ± 0.26	0.130	0.23	−0.07–0.52	97	60.4 ± 20.0	0.026	0.35	0.04–0.65
Often-always	82	55.3 ± 16.7				80	0.58 ± 0.27				75	53.7 ± 18.6			
Duration of pain (days)															
≤7	99	44.8 ± 20.7	0.001	0.50	0.21–0.79	93	0.63 ± 0.28	0.093	0.25	−0.04–0.54	91	59.6 ± 19.3	0.053	0.36	0.05–0.66
>7	90	54.8 ± 19.1				90	0.56 ± 0.28				81	52.7 ± 19.4			
Use of painkillers (number)															
≤2	101	45.6 ± 21.6	0.006	0.40	0.12–0.68	91	0.59 ± 0.29	1.000	0	−0.29–0.29	90	59.0 ± 19.6	0.034	0.32	0.02–0.61
>3	99	53.5 ± 18.0				92	0.59 ± 0.27				92	52.9 ± 18.9			
Infertility															
No	137	46.2 ± 20.0	0.191	0.18	−0.09–0.45	127	0.56 ± 0.30	0.011	−0.36	−0.64– to −0.68	115	54.5 ± 19.9	0.026	−0.32	−0.61 to −0.04
Yes	85	49.9 ± 21.1				84	0.66 ± 0.24				81	60.9 ± 19.4			
Prior surgery for endometriosis															
No	127	47.5 ± 21.2	0.422	0.11	−0.16–0.38	124	0.60 ± 0.28	0.801	0.04	−0.24–0.31	116	58.7 ± 20.0	0.144	0.21	−0.07–0.50
Yes	89	49.8 ± 20.0				83	0.59 ± 0.28				80	54.5 ± 19.2			
ASRM surgical stage															
I–II	56	52.1 ± 21.9	0.160	−0.23	−0.56–0.09	54	0.54 ± 0.29	0.136	−0.25	−0.58–0.08	47	55.0 ± 20.1	0.907	0.02	−0.33–0.37
III–IV	106	47.4 ± 19.2				101	0.61 ± 0.27				98	55.4 ± 18.8			
Type of DIE[c]															

Continued

Table II *Continued*

Known groups	EHP-5					EQ-5D descriptive system					EQ-VAS				
	N[a]	Mean ± SD	*P*-value	Effect size	95% CI effect size	N[a]	Mean ± SD	*P*-value	Effect size	95% CI effect size	N[a]	Mean ± SD	*P*-value	Effect size	95% CI effect size
Anterior DIE															
None	165	48.0 ± 21.6	0.583	0.09	−0.23–0.41	159	0.59 ± 0.29	0.677	−0.07	−0.40–0.26	151	56.1 ± 20.2	0.423	0.14	−0.20–0.48
Bladder	47	49.9 ± 18.3				45	0.61 ± 0.26				42	58.9 ± 19.2			
Posterior DIE															
None-USL	115	48.5 ± 21.5	0.890	0.02	−0.25–0.29	111	0.58 ± 0.28	0.459	−0.11	−0.38–0.17	103	57.3 ± 20.5	0.757	0.05	−0.33–0.24
Vagina-intestine	93	48.9 ± 19.7				89	0.61 ± 0.29				86	56.4 ± 19.1			
Endometrioma															
No	128	48.2 ± 20.9	0.841	0.03	−0.25–0.31	124	0.59 ± 0.28	0.469	0.11	−0.18–0.39	74	58.5 ± 19.2	0.350	0.14	−0.15–0.43
Yes	80	48.8 ± 21.2				76	0.62 ± 0.29				115	55.7 ± 20.6			

USL, Uterosacral. Ligament Comparisons for mean differences were calculated with a two-tailed *t*-test.
Effect sizes were calculated with Cohen's *d* method, $P < 0.05$.
For the EHP-5, lower scores indicate better health status.
[a]Number of participants for whom score could be computed.
[b]Measured by a VAS.
[c]According to Chapron DIE classification (Chapron *et al.*, 2003).

centers for endometriosis surgery and therefore the population of patients was in general consulting after failure of first-line medical treatment. It is therefore important, when interpreting our results, to take into account that the study is mostly based on patients treated surgically. Finally, the dropout rate of 65% leaving only 14 responses in the medical treatment group, results in wide CIs that happen to adversely affect EQ-5D responsiveness with respect to medical treatment. Our aim was to compare the overall responsiveness of the two questionnaires rather than focusing on responsiveness according to treatment subgroups, for which interpretation is limited. It is nonetheless important to underline the fact that the EHP-5 proved to have clear-cut responsiveness in all subgroups studied including the one with medical treatment alone.

An unexpected finding was the fact that EHP-5 and EQ-5D had equivalent responsiveness. It is a known fact that condition-specific scales like EHP-5 are supposed to be more sensitive than generic ones such as EQ-5D. However, in a meta-analysis assessing the responsiveness of generic and specific instruments in randomized controlled trials, it was found that specific instruments were more responsive than generic ones when there was a strong underlying therapeutic effect but not in case of a weak or non-significant effect (Wiebe *et al.*, 2003). This was obviously the case in our study since all patients were included in the analysis, regardless of the type or the completeness of treatment. In addition, it is possible that the type of surgery performed in patients with severe endometriosis, with a long past history of pain, can result in medium therapeutic effect (Vercellini *et al.*, 2009). However, the fact that responsiveness was in general excellent for both scales pleads in favor of our methodology.

The properties of EHP-5 have not been fully studied. Even in its original English version, construct validity and responsiveness were never evaluated (Jones *et al.*, 2004a). The best effect sizes found for EHP-5 were those concerning pain intensity, which were statistically better than those for EQ-5D. Pain being the main symptom of endometriosis disease, this result indicates that EHP-5 reflects the disease's main characteristics accurately, probably thanks to the fact that it was built from face-to-face interviews with patients. Three studies, including the one we previously published, also found that EHP-5 discriminated well between patients according to their symptoms, except for infertility (Goshtasebi *et al.*, 2011; Fauconnier *et al.*, 2013; Selcuk *et al.*, 2015). Unlike ours, none of the studies about EHP-5 evaluated the impact of the extent of the disease on the scale score. As regards to the ASRM stage, the lack of correlation between pain intensity and the extent of the disease could be explained by the fact that symptoms result more from sensory or neural mechanisms related to the depth of the implants than from the spread of the disease as measured by the ASRM classification (Fauconnier *et al.*, 2013).

Although EQ-5D is one of the HrQoL questionnaires most used in endometriosis (Gao *et al.*, 2006), it has not been validated specifically for endometriosis. Here, EQ-5D was sensitive to the presence of painful symptoms. However, these size effects were moderate and the fact that most of the scores tended to be low or inversely correlated to patients' characteristics corroborates the hypothesis that as a generic instrument, it may provide inadequate coverage of certain dimensions of health that are important for patients in given conditions (Lin *et al.*, 2013). However, the good sensitivity to change of the EQ-5D has been used in clinical studies evaluating the impact of surgery on quality of life in endometriosis patients (Garry *et al.*, 2000; Abbott *et al.*,

Table III **Responsiveness of EHP-5 and the EQ-5D: for all patients and those who reported they were better on the Clinical Global Impression-Improvement (CGI-I) scale.**

Groups of patients	EHP-5				EQ-5D descriptive system				EQ-VAS			
	Mean ± SD	*P*-value	Effect size	95% CI effect size	Mean ± SD	*P*-value	Effect size	95% CI effect size	Mean ± SD	*P*-value	Effect size	95% CI effect size
All patients (*N*[a])	*N* = 132	<0.001	0.81	0.56–1.56	*N* = 127	<0.001	0.95	0.68–1.20	*N* = 120	<0.001	0.85	0.58–1.11
Baseline	48.1 ± 20.8				0.59 ± 0.28				56.3 ± 20.5			
Follow-up	29.6 ± 24.6				0.82 ± 0.20				73.4 ± 19.7			
All patients with surgical treatment (*N*[a])	*N* = 119	<0.001	0.80	0.53–1.06	*N* = 113	<0.001	0.95	0.67–1.22	*N* = 108	<0.001	0.92	0.63–1.19
Baseline	48.7 ± 21.0				0.59 ± 0.27				54.9 ± 20.3			
Follow-up	30.3 ± 24.8				0.82 ± 0.21				73.3 ± 19.9			
Complete surgery (*N*)	*N* = 104	<0.001	0.81	0.53–1.09	*N* = 100	<0.001	0.91	0.61–1.21	*N* = 94	<0.001	0.95	0.65–1.25
Baseline	49.1 ± 21.4				0.58 ± 0.28				54.4 ± 19.6			
Follow-up	30.1 ± 25.3				0.82 ± 0.21				73.1 ± 19.6			
Incomplete surgery (*N*)	*N* = 15	0.057	0.72	−0.03–1.44	*N* = 13	0.009	1.12	0.26–1.90	*N* = 14	0.077	0.70	−0.08–1.44
Baseline	46.3 ± 18.7				0.68 ± 0.20				57.7 ± 25.6			
Follow-up	31.4 ± 22.3				0.86 ± 0.11				74.7 ± 23.1			
All patients with medical treatment (*N*[a])	*N* = 13	0.027	0.93	0.07–1.70	*N* = 14	0.066	0.73	−0.06–1.47	*N* = 12	0.493	0.28	−0.53–1.08
Baseline	42.3 ± 18.2				0.63 ± 0.34				69.1 ± 18.5			
Follow-up	23.4 ± 22.4				0.83 ± 0.19				74.2 ± 17.3			
Patients who reported they were better (*N*[a])	*N* = 118	<0.001	0.94	0.67–1.21	*N* = 114	<0.001	1.03	0.75–1.30	*N* = 106	<0.001	0.98	0.72–1.31
Baseline	48.0 ± 20.2				0.59 ± 0.27				55.5 ± 20.1			
Follow-up	27.6 ± 22.9				0.83 ± 0.19				74.6 ± 18.9			
Patients with surgical treatment who reported they were better (*N*[a])	*N* = 106	<0.001	0.92	0.63–1.20	*N* = 101	<0.001	1.11	0.81–1.41	*N* = 95	<0.001	1.05	0.80–1.42
Baseline	48.2 ± 20.7				0.58 ± 0.27				54.2 ± 20.2			
Follow-up	28.1 ± 23.0				0.84 ± 0.19				74.9 ± 19.3			
Complete surgery (*N*)	*N* = 94	<0.001	0.92	0.61–1.22	*N* = 91	<0.001	1.11	0.79–1.43	*N* = 84	<0.001	1.07	0.74–1.39
Baseline	48.8 ± 20.8				0.57 ± 0.27				54.4 ± 19.4			
Follow-up	28.4 ± 23.5				0.83 ± 0.19				74.7 ± 18.6			
Incomplete surgery (*N*)	*N* = 12	0.033	0.93	0.06–1.73	*N* = 10	0.012	1.25	0.25–2.15	*N* = 11	0.032	0.91	0–1.75
Baseline	43.6 ± 20.1				0.68 ± 0.19				52.8 ± 26.9			
Follow-up	25.3 ± 19.4				0.87 ± 0.10				76.4 ± 25.0			
	N = 12	0.014	1.09	0.18–1.81	*N* = 13	0.064	0.76	−0.06–1.53	*N* = 11	0.437	0.34	−0.52–1.16

Continued

Table III *Continued*

Groups of patients	EHP-5				EQ-5D descriptive system				EQ-VAS			
	Mean ± SD	*P*-value	Effect size	95% CI effect size	Mean ± SD	*P*-value	Effect size	95% CI effect size	Mean ± SD	*P*-value	Effect size	95% CI effect size
Patients with medical treatment who reported they were better (N^a)												
Baseline	45.9 ± 15.3				0.60 ± 0.34				66.3 ± 16.5			
Follow-up	24.9 ± 22.6				0.81 ± 0.19				71.8 ± 16.0			

Comparison for mean differences were calculated with a two-tailed *t*-test.
Effect sizes were calculated with Cohen's *d* method, $P < 0.05$.
For the EHP-5, lower scores indicate better health status.
[a]Number of participants for whom score could be computed.

2004). Our study confirms the fact that EQ-5D presents excellent responsiveness in case of surgical therapy and is therefore recommended in clinical research focusing on surgery.

The other questionnaires used most in endometriosis are SF-36 and EHP-30, (Gao *et al.*, 2006). SF-36 has been specifically validated for endometriosis, while EHP-30 is specific to endometriosis (Jones *et al.*, 2001; Stull *et al.*, 2014). EHP-30 has been recommended by several medical societies (Vincent *et al.*, 2010; Khong *et al.*, 2010). However, each of these two scales is quite long to complete and is made up of several subscales with a score for each subscale. Consequently, when evaluating responsiveness, that of EHP-30 was determined item by item and that for SF-36 subscale by subscale (Jones *et al.*, 2004b; Stull *et al.*, 2014). This may be very impractical for several reasons. First, in case of daily use, for it has been reported that the measurement of health was often limited by the 'time of use' and the multiple scores may create a problem for interpretation by practitioners (Coste *et al.*, 1997). Next, data quality can be affected when instruments that collect large amounts of information are used (Doward *et al.*, 2004).

PRO instruments are now those recommended by the Food and Drug Agency (Turk *et al.*, 2006) and by the consensus of an international meeting on chronic pain and clinical trials (Acquadro *et al.*, 2003) for assessing the impact of the disease or the treatment administered. From this point of view, the clinimetric properties that we demonstrated for both EHP-5 and EQ-5D may be of great interest for research and policy setting, each for a specific use. Both EHP-5 and EQ-D have already demonstrated their effectiveness as PRO in research trials evaluating surgical treatment of endometriosis (Minas and Dada, 2014; De la Hera-Lazaro *et al.*, 2016; Garry *et al.*, 2000; Abbott *et al.*, 2004). Responsiveness in case of medical treatment has yet to be confirmed by larger studies. EQ-5D, as a generic instrument, is particularly useful for epidemiological studies since it allows comparison of managements or populations regardless of the disease (Patrick and Deyo, 1989). Next, as a cost-utility instrument, the 'health states' obtained by the responses to the EQ-5D descriptive system lead to the calculation of quality-adjusted life years, particularly useful for economic evaluation of health-care interventions (Whitehead and Ali, 2010).

As regards clinical practice, routine evaluation of HrQOL in women who suffer from endometriosis is of great importance both for the health-care provider and the patient (Higginson and Carr, 2001). With this in mind, EHP-5 appears to be a better candidate than EQ-5D. Indeed, it is simpler and easier to interpret, facilitating evaluation of the baseline quality of life. Its structure permits screening for hidden problems like social ones for example (Higginson and Carr, 2001). Next, since EQ-5D proved to have lower construct validity, it may not be as valuable as EHP-5 to help decision-making at T_0 (i.e. time of diagnosis). Therefore, EHP-5 may contribute to improve the provider-initiated communication and the shared decision-making process with the patient, which is particularly valuable since in endometriosis the treatment focuses on the patient rather than the disease (Dunselman *et al.*, 2014). The use of a good instrument, such as EHP-5, is a key element that could identify factors in determining the most effective and personalized treatment. From this point of view, a perspective for further studies could be to explore the prognostic value of EHP-5 before surgery to define a threshold above which surgery would be of benefit for the patient.

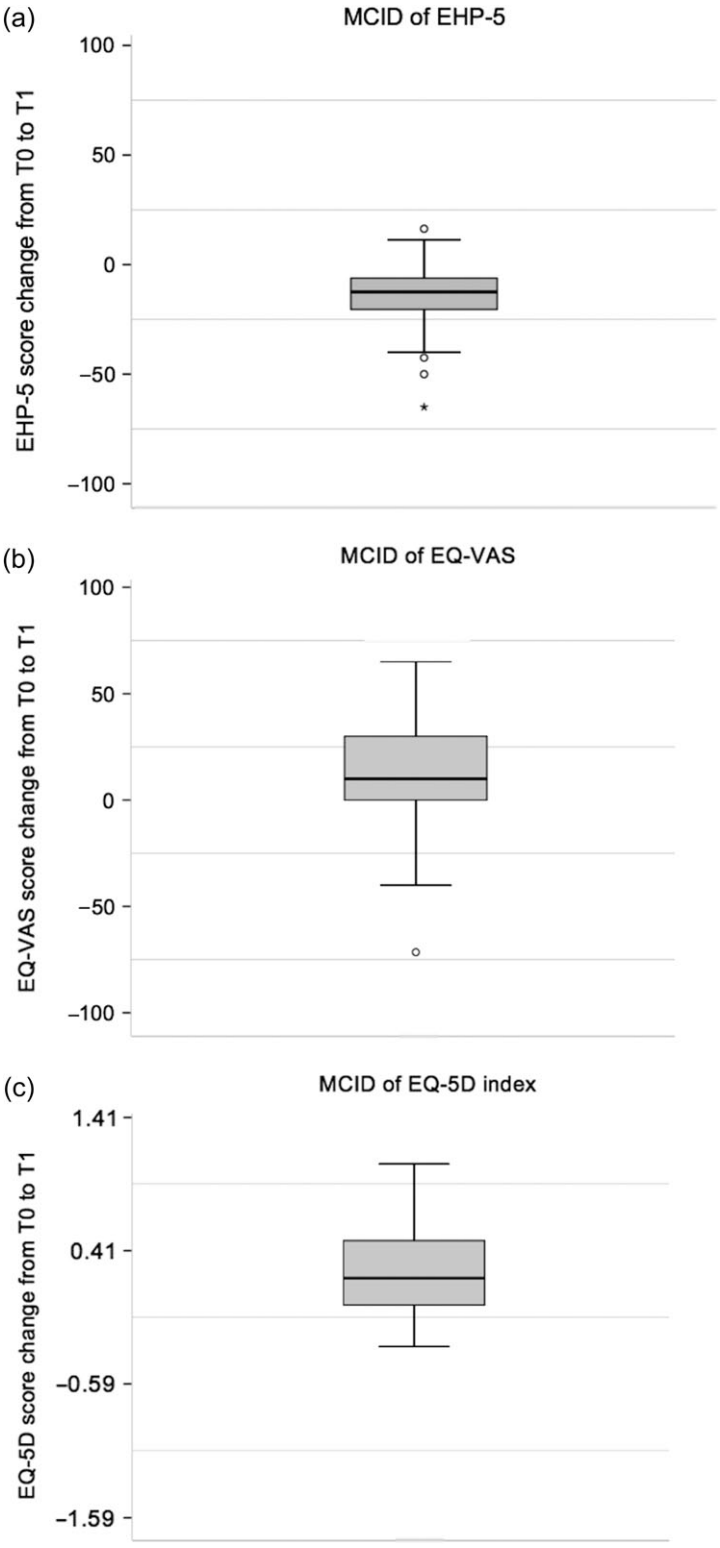

Figure 2 Box plots of the scale score change from T_0 to T_1, in the subgroups of patients who reported they were 'somewhat better' or 'better' after treatment, according to the Clinical Global Impression-Improvement (CGI-I) scale. (**a**) Endometriosis Health Profile-5 (EHP-5), (**b**) EuroQol Visual Analogic Scale (EQ-VAS), (**c**) EQ-5D, MCID, minimal clinically important difference. Notes: number of patients = 37. The mean score change corresponds to the MCID. The open circles represent aberrant values of patients treated with medical treatment; the star represents the aberrant value of a patient treated with surgery.

Acknowledgements

The Direction à la Recherche Clinique et à l'Innovation of Versailles, France for sponsoring the study, and for its help in data collection. We are also grateful to Georgina Jones, creator of the EHP-5, for her availability and for giving us the EHP-5 user manual.

Authors' roles

A.F. was coordinator of the project. A.F., P.P. and G.T. conceived and designed the study. P.P., A.F., G.A. and G.T. contributed to data collection and/or performed surgical procedures and medical follow-up. G.A., A.F. and C.H. analyzed and interpreted the data. G.A. and A.F. wrote the manuscript. All authors approved the final submitted version of the manuscript.

Funding

The Direction à la Recherche Clinique et à l'Innovation of Versailles.

Conflict of interest

None declared.

References

Abbott J, Hawe J, Hunter D, Holmes M, Finn P, Garry R. Laparoscopic excision of endometriosis: a randomized, placebo-controlled trial. *Fertil Steril* 2004;**4**:878–884.

Acquadro C, Berzon R, Dubois D, Leidy NK, Marquis P, Revicki D, Rothman M. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the patient-reported outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health* 2003;**5**:522–531.

ASRM. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertil Steril* 1997;**5**:817–821.

Bazot M, Lafont C, Rouzier R, Roseau G, Thomassin-Naggara I, Daraï E. Diagnostic accuracy of physical examination, transvaginal sonography, rectal endoscopic sonography and magnetic resonance imaging to diagnose deep infiltrating endometriosis. *Fert Steril.* 2009;**92**:1825–1833.

Bazot M, Bornier C, Dubernard G, Roseau G, Cortez A, Darai E. Accuracy of magnetic resonance imaging and rectal endoscopic sonography for the prediction of location of deep pelvic endometriosis. *Hum Reprod* 2007;**5**:1457–1463.

Canis M. Laparoscopic surgery for deep endometriosis. About 118 cases. *Gynecol Obstet Fertil* 2007;**1**:77–78.

Chapron C, Fauconnier A, Vieira M, Barakat H, Dousset B, Pansini V, Vacher-Lavenu MC, Dubuisson JB. Anatomical distribution of deeply infiltrating endometriosis: surgical implications and proposition for a classification. *Hum Reprod* 2003;**1**:157–161.

Chevalier J, de Pouvourville G. Valuing EQ-5D using time trade-off in France. *Eur J Health Econ* 2013;**1**:57–66.

Coste J, Guillemin F, Pouchot J, Fermanian J. Methodological approaches to shortening composite measurement scales. *J Clin Epidemiol* 1997;**3**:247–252.

Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;**5**:395–407.

De Graaff AA, D'Hooghe TM, Dunselman GA, Dirksen CD, Hummelshoj L, Simoens S. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. *Hum Reprod* 2013;**10**:2677–2685.

De la Hera-Lazaro CM, Munoz-Gonzalez JL, Perez RO, Vellido-Cotelo R, Diez-Alvarez A, Munoz-Hernando L, Alvarez-Conejo C, Jimenez-Lopez JS. Radical surgery for endometriosis: analysis of quality of life and surgical procedure. *Clin Med Insights Womens Health* 2016;**9**:7–11.

Doward LC, Meads DM, Thorsen H. Requirements for quality of life instruments in clinical research. *Value Health* 2004;**7**:S13–S16.

Dubernard G, Piketty M, Rouzier R, Houry S, Bazot M, Darai E. Quality of life after laparoscopic colorectal resection for endometriosis. *Hum Reprod* 2006;**5**:1243–1247.

Dunselman GA, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, De Bie B, Heikinheimo O, Horne AW, Kiesel L, Nap A *et al*. ESHRE guideline: management of women with endometriosis. *Hum Reprod* 2014;**3**:400–412.

Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, Haythornthwaite JA, Jensen MP, Kerns RD, Ader DN *et al*. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;**2**:105–121.

EuroQol Group. EuroQol Group-a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**3**:199–208.

Fauconnier A, Huchon C, Chaillou L, Aubry G, Renouvel F, Panel P. Development of a French version of the Endometriosis Health Profile 5 (EHP-5): cross-cultural adaptation and psychometric evaluation. *Qual Life Res* 2017;**1**:213–220.

Fauconnier A, Staraci S, Huchon C, Roman H, Panel P, Descamps P. Comparison of patient- and physician-based descriptions of symptoms of endometriosis: a qualitative study. *Hum Reprod* 2013;**10**:2686–2694.

Fourquet J, Baez L, Figueroa M, Iriarte RI, Flores I. Quantification of the impact of endometriosis symptoms on health-related quality of life and work productivity. *Fertil Steril* 2011;**1**:107–112.

Gao X, Yeh YC, Outley J, Simon J, Botteman M, Spalding J. Health-related quality of life burden of women with endometriosis: a literature review. *Curr Med Res Opin* 2006;**9**:1787–1797.

Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002;**7351**:1417.

Garry R, Clayton R, Hawe J. The effect of endometriosis and its radical laparoscopic excision on quality of life indicators. *Bjog* 2000;**1**:44–54.

Gerlinger C, Schumacher U, Faustmann T, Colligs A, Schmitz H, Seitz C. Defining a minimal clinically important difference for endometriosis-associated pelvic pain measured on a visual analog scale: analyses of two placebo-controlled, randomized trials. *Health Qual Life Outcomes* 2010;**8**:138.

Giudice LC. Clinical practice: endometriosis. *N Engl J Med* 2010;**25**:2389–2398.

Goshtasebi A, Nematollahzadeh M, Hariri FZ, Montazeri A. The Short Form Endometriosis Health Profile (EHP-5): translation and validation study of the Iranian version. *J Ovarian Res* 2011;**4**:11.

Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;**8**:622–629.

Higginson IJ, Carr AJ. Using quality of life measures in clinical setting. *Bmj* 2001;**7297**:1297–1300.

Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;**4**:407–415.

Jones G, Jenkinson C, Kennedy S. Development of the Short Form Endometriosis Health Profile Questionnaire: the EHP-5. *Qual Life Res* 2004a;**3**:695–704.

Jones G, Jenkinson C, Kennedy S. Evaluating the responsiveness of the Endometriosis Health Profile Questionnaire: the EHP-30. *Qual Life Res* 2004b;**3**:705–713.

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;**2**:258–264.

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;**2**:501–511.

Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;**3**:S178–S189.

Khong SY, Lam A, Luscombe G. Is the 30-item Endometriosis Health Profile (EHP-30) suitable as a self-report health status instrument for clinical trials? *Fertil Steril* 2010;**5**:1928–1932.

Lin FJ, Longworth L, Pickard AS. Evaluation of content on EQ-5D as compared to disease-specific utility measures. *Qual Life Res* 2013;**4**:853–874.

Minas V, Dada T. Laparoscopic treatment of endometriosis and effects on quality of life: a retrospective study using the short form EHP-5 endometriosis specific questionnaire. *J Obstet Gynaecol* 2014;**4**:336–340.

Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, Jenkinson C, Kennedy SH, Zondervan KT. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertil Steril* 2011;**2**:366–373.e368.

Patrick DL, Deyo RA. Generic and disease-specific measure in assessing health status and quality of life. *Med Care* 1989;**3**:S217–S232.

Renouvel F, Fauconnier A, Pilkington H, Panel P. Linguistic adaptation of the Endometriosis Health Profile 5: EHP 5. *J Gynecol Obstet Biol Reprod (Paris)* 2009;**5**:404–410.

Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;**2**:102–109.

Schliep KC, Mumford SL, Peterson CM, Chen Z, Johnstone EB, Sharp HT, Stanford JB, Hammoud AO, Sun L, Buck Louis GM. Pain typology and incident endometriosis. *Hum Reprod* 2015;**10**:2427–2438.

Selcuk S, Sahin S, Demirci O, Aksoy B, Eroglu M, Ay P, Cam C. Translation and validation of the Endometriosis Health Profile (EHP-5) in patients with laparoscopically diagnosed endometriosis. *Eur J Obstet Gynecol Reprod Biol* 2015;**185**:41–44.

Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, Brodszky V, Canis M, Colombo GL, DeLeire T *et al*. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum Reprod* 2012;**5**:1292–1299.

Stull DE, Wasiak R, Kreif N, Raluy M, Colligs A, Seitz C, Gerlinger C. Validation of the SF-36 in patients with endometriosis. *Qual Life Res* 2014;**1**:103–117.

Turk DC, Dworkin RH, Burke LB, Gershon R, Rothman M, Scott J, Allen RR, Atkinson JH, Chandler J, Cleeland C *et al*. Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. *Pain* 2006;**3**:208–215.

Vercellini P, Crosignagni PG, Abbiati A, Somigliana E, Vigano P, Fedele L. The effect of surgery for symptomatic endometriosis: the other side of the story P. *Hum Reprod Update* 2009;**2**:177–188.

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;**2**:299–304.

Vincent K, Kennedy S, Stratton P. Pain scoring in endometriosis: entry criteria and outcome measures for clinical trials. Report from the Art and Science of Endometriosis meeting. *Fertil Steril* 2010;**1**:62–67.

Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull* 2010;**1**:5–21.

Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol* 2003;**1**:52–60.