Health-related quality of life in women with endometriosis

psychometric validation of the Endometriosis Health Profile 30 questionnaire using confirmatory factor analysis

Hansen, K E; Lambek, R; Røssaak, K; Egekvist, A G; Marschall, H; Forman, A; Kesmodel, U S

Published in:
Human reproduction open

DOI (link to publication from Publisher):
10.1093/hropen/hoab042

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Human Reproduction Open, pp. 1–11, 2021
https://doi.org/10.1093/hropen/hoab042

human reproduction open

ORIGINAL ARTICLE

Health-related quality of life in women with endometriosis: psychometric validation of the Endometriosis Health Profile 30 questionnaire using confirmatory factor analysis

K.E. Hansen [1,*], R. Lambek[2], K. Røssaak[3], A.G. Egekvist[4,5], H. Marschall[2], A. Forman[4,5], and U.S. Kesmodel[6,7]

[1]Department of Public Health, Aarhus University, Aarhus, Denmark [2]Department of Psychology and Behavioural Sciences, School of Business and Social Sciences, Aarhus University, Aarhus, Denmark [3]Department of Obstetrics and Gynaecology, Rigshospitalet, Copenhagen, Denmark [4]Department of Obstetrics and Gynaecology, Aarhus University Hospital, Aarhus, Denmark [5]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark [6]Department of Obstetrics and Gynaecology, Aalborg University Hospital, Aalborg, Denmark [7]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

*Correspondence address. Department of Public Health, Aarhus University, 8000 Aarhus, Denmark. Tel: +45-87150000; E-mail: keh@ph.au.dk https://orcid.org/0000-0003-0243-7399

Submitted on May 24, 2021; resubmitted on October 27, 2021; editorial decision on November 02, 2021

STUDY QUESTION: Which of the competing models of the Endometriosis Health Profile 30 Questionnaire (EHP-30) factor structure is best supported by confirmatory factor analysis (CFA)?

SUMMARY ANSWER: Findings support a five-factor first-order model of the EHP-30, thereby lending support to the model originally suggested by the questionnaire developers.

WHAT IS KNOWN ALREADY: Endometriosis has a negative impact on quality of life, and measures specifically developed to address this impact, such as the EHP-30, are vital in research and disease management. Previous studies have found different models of the EHP-30 factor structure, and generated uncertainty regarding how to use the questionnaire. CFA can be applied to compare competing factor models and determine the underlying structure of a questionnaire.

STUDY DESIGN, SIZE, DURATION: This cross-sectional multicenter study included 304 women with endometriosis recruited from three different public health service endometriosis clinics (referral centers for treatment of severe endometriosis) and the Danish Endometriosis Patients Association from 2014 to 2015.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Diagnosis of endometriosis was confirmed in medical records for 84.2% and by histology for 66.8% of participants. Questionnaires (the licensed Danish version of the EHP-30) were sent by post two times with a 6- to 12-week interval. CFA was used to examine construct validity and Bland–Altman plots to examine test–retest reliability and the convergent validity with the Short Form 36 version 2.

MAIN RESULTS AND THE ROLE OF CHANCE: Response rate was high (87.6%). CFA supported the original first-order five-factor structure of the EHP-30, and thereby, the use of five separate scale-scores in clinical and research practice. Visual inspection of Bland–Altman plots suggested excellent test–retest reliability of the EHP-30 and supported the use of a disease specific quality of life instrument for women with endometriosis.

LIMITATIONS, REASONS FOR CAUTION: Diagnosis could not be confirmed through histology data in 33.2% of participants. However, subgroup analyses based on women with confirmed histology only, yielded similar results. Data related to menstrual cycle stage and the use of hormonal and pain medication during questionnaire completion were not collected. A larger study, including data from different countries on different continents, would be better designed to exclude potential population bias.

WIDER IMPLICATIONS OF THE FINDINGS: EHP-30, with its original five-factor structure, appears to be a valid, stable, and specific quality of life measure for women with endometriosis. It seems easy to understand, quick to administer, and importantly, scoring might be unaffected by cyclical/menstrual pain symptoms related to endometriosis. The finding of a five-factor model from different studies across several countries supports the crosscultural validity of the EHP-30.

STUDY FUNDING/COMPETING INTEREST(S): This work was supported by the Danish Endometriosis Association, which is a non-governmental organization run by women with endometriosis and by a scholarship from the Health Research Fund of Central Denmark Region. The authors have no conflicts of interest.

TRIAL REGISTRATION NUMBER: The Danish Data Protection Agency (J.nr: 2013-41-2264).

Key words: women / reproduction / health-related quality of life / psychometric properties / endometriosis / the Endometriosis Health Profile 30 questionnaire / confirmatory factor analysis (CFA) / menstrual stage / cross cultural validity / Bland–Altman plots

WHAT DOES THIS MEAN FOR PATIENTS?

The Endometriosis Health Profile (EHP-30) is a questionnaire developed to measure the extent to which the disease of endometriosis has a negative impact on quality of life in patients and is often used to help understand the effect of different endometriosis treatments. The questionnaire contains 30 questions that women with endometriosis themselves find important to their quality of life. The questions can be divided into five subscales covering 'pain', 'control and powerlessness', 'social support', 'emotional wellbeing' and 'self-image'. The questionnaire has been translated from English into different languages in order to be used in different countries and cultures. However, studies examining the measurement quality of these translated versions have indicated that a different number of subscales is applicable in different countries. This has generated uncertainty about how to use the questionnaire. This study was carried out using more sophisticated statistical methods than previously in order to examine the number of subscales to be used when applying the questionnaire in research and clinical practice. The results of the study support the use of the original five EHP-30 subscales and demonstrate that the questionnaire measures a unique impact of endometriosis on health-related quality of life not covered by more general quality of life questionnaires.

Introduction

Endometriosis is a common chronic gynecological disease with an estimated prevalence of 5–10% among women of reproductive age (Ferrero *et al.*, 2010). The disorder is defined by the presence of endometrium-like tissue outside the uterus in the abdominal cavity that causes inflammation and adhesions (Vigano *et al.*, 2004). Long-term symptoms include cyclic and chronic pelvic pain, dyspareunia, dyschezia, dysuria, abnormal bleeding, infertility and fatigue (Hansen *et al.*, 2014), but the disease is also associated with severe psychological distress (Culley *et al.*, 2013; De Graaff *et al.*, 2013; Chen *et al.*, 2016) and a significant negative impact on health-related quality of life (HRQoL) (Marinho *et al.*, 2018). Therefore, assessment and incorporation of HRQoL into research and treatment of endometriosis have become critical for disorder management (NICE guideline 73).

Measuring HRQoL by use of a generic questionnaire, such as the Short Form 36 (SF-36) (Garratt *et al.*, 1993; Bjorner *et al.*, 1998a,b), does not provide information about the potentially unique impact of a specific disease on HRQoL. Therefore, in order to measure the specific impact of endometriosis on HRQoL, the Endometriosis Health Profile 30 Questionnaire (EHP-30) was developed (Jones *et al.*, 2001). The original English version of the EHP-30 was found to have a high degree of internal consistency and test–retest reliability (Jones *et al.*, 2001), and was found to be more sensitive to change than the generic SF-36 in women undergoing treatment for endometriosis-associated pain (Jones *et al.*, 2004). A recent systematic review indicated that EHP-30 is not only the most commonly used endometriosis-specific questionnaire in the literature but also the most thoroughly validated and reliable questionnaire for measuring HRQoL in women with endometriosis (Bourdel *et al.*, 2019). However, previous studies have found different models of the underlying EHP-30 factor structure and generated uncertainty regarding how to use the questionnaire in clinical and research practice. To date, four studies have supported the original five-factor structure of the EHP-30 (Jenkinson *et al.*, 2008 (American); Jia *et al.*, 2013 (Chinese); Nojomi *et al.*, 2011 (Persian); van de Burgt *et al.*, 2011 (Dutch)), one study found a four-factor structure (Jones *et al.*, 2006 (English)), and one obtained mixed results and suggested four to five factors (Chauvet *et al.*, 2017 (French)). Two recent studies have suggested a three-factor structure, however, the statistical analyses and results in these studies were unclear (Grundström *et al.*, 2020 (Swedish); Verket *et al.*, 2018 (Norwegian)). While previous studies have been informative, they have all used exploratory factor analysis (i.e. principal component analysis, PCA) to evaluate the factor structure of the EHP-30. Consequently, further empirical investigation with the use of more sophisticated statistical techniques, such as CFA, is warranted. This method will also allow a direct comparison of the different factor models found in previous research in order to determine the best model.

Hence, the aim of the present study was, for the first time, to use CFA to compare competing factor models identified in previous research in order to examine the latent factor structure of the EHP-30 questionnaire.

Materials and methods

Study design and participants

This cross-sectional multicenter study was conducted from 2014 to 2015. Patients (18 years and above) with endometriosis were recruited when they attended for checkup of disease development and the effect of treatment at one of two referral centers for the treatment of severe endometriosis in the Danish public health service (Aarhus University Hospital and Copenhagen University Hospital), from an outpatient endometriosis clinic at Svendborg Hospital, and from the Danish Endometriosis Patients Association. Participants gave written informed consent and received the questionnaires by postal mail.

Endometriosis diagnosis and histology

Diagnosis and histology were verified through patient records where possible. Access to patients' medical records outside of Aarhus and Copenhagen University Hospitals was not granted by the authorities. All patients were retained in the analyses irrespective of confirmed diagnosis. The primary analyses were rerun post hoc including only the subgroup of patients with histologically confirmed endometriosis.

Questionnaires

EHP-30 is a patient-generated and disease specific self-report questionnaire containing aspects that women with endometriosis themselves find important to their QoL. It consists of 30 items to which respondents can choose between the answers: Never (0); Rarely (1), Sometimes (2); Often (3); and Always (4). The 30 items are divided into five subscales covering 'pain', 'control and powerlessness', 'social support', 'emotional wellbeing' and 'self-image'. Each scale is standardized on a scale from 0 to 100, with lower scores indicating better QoL (Jones *et al.*, 2001). We used the licensed Danish version of the EHP-30, which has been linguistically validated according to standard guidelines (Beaton *et al.*, 2000) and included forward translation by two independent native speaking translators, reconciliation of the two into a third version, backtranslation by two independent native speaking translators and review against the original English version. Furthermore, the Danish version underwent clinical review and cognitive debriefing by a specialized clinician and five Danish patients with endometriosis. The EHP-30 questionnaire was sent two times by postal mail to participants with an interval of more than 4 weeks (i.e. more than a menstrual period) (Range: 6–12 weeks). Along with the first EHP-30 questionnaire, patients also received a short background questionnaire and the SF-36v2 questionnaire, which measures general well-being and health-related QoL (Garratt *et al.*, 1993). Patients also answered questions concerning pain and other endometriosis-related symptoms.

Statistical analysis

First, scores on the EHP-30 were examined with respect to variability and homogeneity (Gorecki *et al.*, 2013). This was done by determining the amount of missing data at item level and examining the score distribution, restricted range, floor and ceiling effects, as well as item means and SDs.

Second, the construct validity of scores on the EHP-30 was examined using CFA. Five-factor analytic models were specified and estimated (Fig. 1). Model 1 specified one latent variable on which all 30

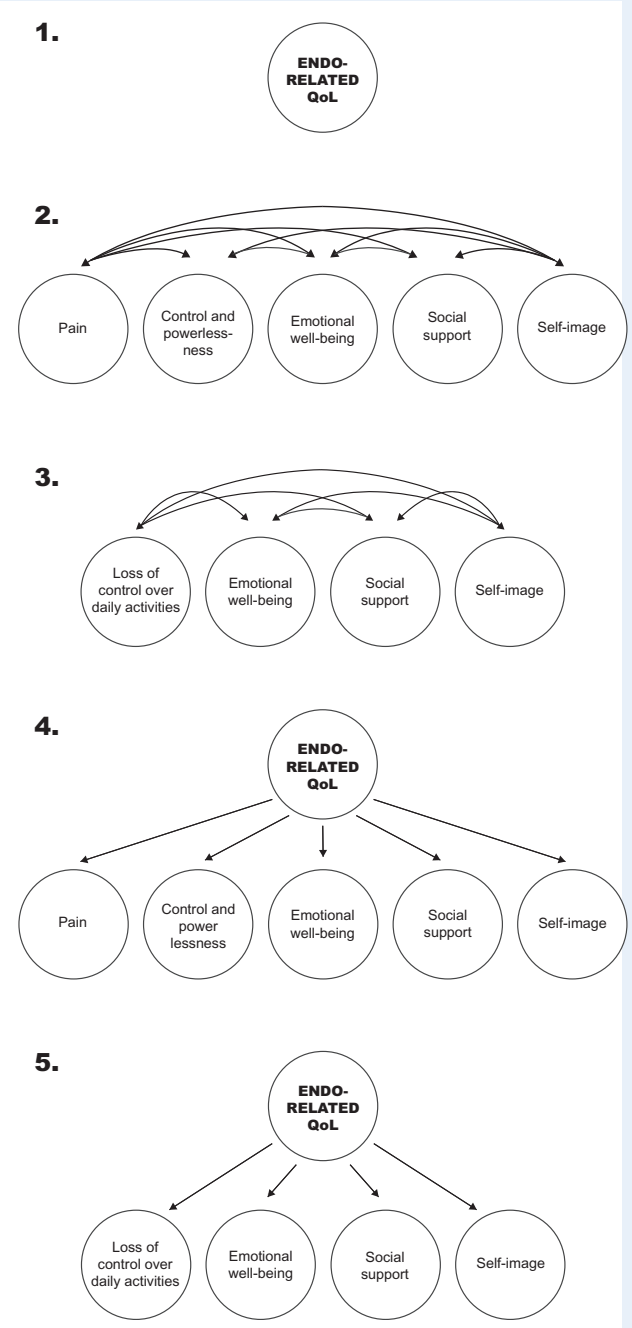

Figure 1. The five EHP-30 factor models compared by CFA. Model 1 specifies one latent variable on which all 30 items load. Model 2 specifies five correlated latent variables (Pain, Control and powerlessness, Emotional well-being, Social support and Self-image). Model 3 specifies four latent variables (Loss of control over daily activities, Emotional well-being, Social support and Self-image). Model 4 is a higher-order variant of Model 2. Model 5 is a higher-order variant of Model 3. EHP-30, Endometriosis Health Profile 30 Questionnaire; Endo-related QoL, endometriosis-related quality of life.

items loaded, and was labeled 'endometriosis-related quality of life'; this represents the single-factor solution as proposed by Jenkinson et al. (2008). Model 2 specified five correlated latent variables (Pain, Control and powerlessness, Emotional well-being, Social support and Self-image) and reflected the original endometriosis-related QoL model suggested by Jones et al. (2001). Model 3 specified four latent variables. In this model, three latent variables from Model 2 were retained (Emotional well-being, Social support and Self-image), but Pain and Control and powerlessness were collapsed into one latent variable labeled 'Loss of control over daily activities'. This model was based on the results by Jones et al. (2006) and Chauvet et al. (2017). Higher-order variants of Models 2 and 3 were also specified. This included a single second-order latent variable labeled 'Endometriosis-related quality of life'. These models allowed us to examine whether the variation and covariation among the first-order latent variables could be explained by a single second-order latent variable. For all first-order models the latent variables were correlated, and for all models, measurement error variances were uncorrelated.

Model parameters were estimated using robust maximum likelihood using all available data. Global model fit was based on established guidelines (Jackson et al., 2009) and determined as follows: Chi-square and model degrees of freedom, with a nonsignificant chi-square indicating acceptable model fit; comparative fit index (CFI), and Tucker–Lewis Index (TLI) values >0.95 suggesting excellent model fit and values between 0.90 and 0.95 suggesting acceptable fit; root mean square error of approximation (RMSEA) values of 0.05 or less suggesting good fit and values between 0.05 and 0.08 suggesting reasonable approximate fit; standardized root mean square residual (SRMR) values <0.05 indicating good fit. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used to compare competing models with lower values indicating better fit. The modification indices (MI) were examined to identify additional parameters that, if added to the model, would significantly improve model fit. A strong theoretical rationale for adding additional parameters was required. Finally, difference testing (Satorra and Bentler, 2001) was used to compare the relative fit of competing models (with $P < 0.05$ suggesting that the least restrictive model should be retained).

Internal consistency was examined with Cronbach's alpha. Test–retest reliability was examined using Bland–Altman plots of agreement between measures (Bland and Altman, 1986). Percentages of participants scoring the same value or the same value ±1 between first and second measurement on the EHP-30 questionnaire were calculated.

Finally, the agreement between EHP-30 subscales and relevant SF-36 subscales was examined using Bland–Altman plots, as a measure of convergent validity. Correlations between EHP-30 subscales and an EHP-30 total sum with relevant SF-36 domains were calculated using Pearson's correlation.

The CFAs were estimated using Mplus version 8.1 (Muthén and Muthén, 2017) and the remaining analyses were performed using IBM SPSS Statistics for Windows, version 24 (IBM Corp., Armonk, NY, USA).

Ethical approval

The project was approved by the Danish Data Protection Agency (J.nr: 2013-41-2264). All participants gave written informed consent.

Results

A total of 348 questionnaires were sent out in the first round and 305 were returned (response rate = 87.6%). Subsequently, one

Table 1 Demographic characteristics and pain level of participants.

Characteristics	Participants (N = 298)[a] N (%)	Mean (SD)
Age (range 18–58 years)		35.49 (6.924)
Marital status		
Married/living together	233 (76.6%)	
Single/living alone	46 (15.1%)	
Other	19 (6.3%)	
Children		
Biological		
No children	162 (53.3%)	
1	56 (18.4%)	
2	68 (22.4%)	
3	10 (3.3%)	
>3	2 (0.6%)	
Adopted		
No children	259 (85.2%)	
1	20 (6.6%)	
2	13 (4.3%)	
3	4 (1.3%)	
>3	2 (0.6%)	
Occupation		
Full time	143 (47%)	
Part time	53 (17.4%)	
Freelance/consultant	4 (1.3%)	
Enrolled in education	22 (7.2%)	
Maternity leave	6 (2.0%)	
Flexijob or rehabilitation	24 (8.0%)	
Sick leave, government benefits, or no income	41 (13.5%)	
Other	5 (1.6%)	
Chronic pelvic pain (NRS)		4.102 (2.732)
0	46 (15.1%)	
1	14 (4.6%)	
2	25 (8.2%)	
3	33 (10.9%)	
4	37 (12.2%)	
5	34 (11.2%)	
6	31 (10.2%)	
7	29 (9.5%)	
8	23 (7.6%)	
9	9 (3.0%)	
10	3 (1.0%)	

[a]Missing data (n = 6).
NRS, numeric rating scale.

participant was excluded from the study owing to young age (<18 years), yielding 304 participants in the final sample. Diagnosis was confirmed through patient records in 84.2% (N = 256) of the patients and with confirmed histology in 66.8% (N = 203) of the patients. Mean time since diagnosis was 7.14 years (SD = 5.63, range = 0–25 years) and mean time from symptom onset to diagnosis was 8.26 years (SD = 6.97, range = 0–36 years). A total of 262 participants returned the second questionnaire in round two with 78.2% (N = 205) reporting no treatment changes since answering the first questionnaire. Participants' demographic characteristics and pain levels are described in Table I.

Acceptability and data quality

Item-level missing data, score distributions, means and SDs are presented in Supplementary Table SI. There were few missing data (<0.5%) and only two items (Items 3 and 23) were characterized by restricted range (i.e. no participants used response Category 4 'always'). Scores on Item 23 were also highly skewed (>+1), suggesting some reduced ability to produce score variability. There were no floor or ceiling effects. Overall, this suggests acceptable data quality.

Construct validity

The fit statistics from the CFAs are reported in Table II. The first-order five-factor model (Model 2) and the higher-order five-factor model (Model 4) fitted the data better than the other models. These two models had marginally acceptable fit with CFI/TLI values just below 0.90, RMSEA values <0.08, SRMR values around 0.05, and the lowest AIC/BIC values. Inspection of MI suggested the presence of a large residual covariance between Items 22 and 21 in both models. When that parameter was added, model fit improved substantially (first-order model: $\chi^2 = 1032.634$, df = 394, $P = 0.000$; CFI = 0.911; TLI = 0.902; RMSEA = 0.073, 90% CI = [0.068–0.078]; SRMR = 0.049; AIC = 20821.659; BIC = 21197.079; higher-order model: $\chi^2 = 1056.435$, df = 399, $P = 0.000$; CFI = 0.908; TLI = 0.900; RMSEA = 0.074, 90% CI = [0.068–0.079]; SRMR = 0.052; AIC = 20839.548; BIC = 21196.382). Fit statistics for the two models were similar, albeit difference testing favored the less restrictive model (i.e. the first-order five-factor model) as the better model of the two (TRd = 24.467, df = 5, $P < 0.000$). Factor loadings and factor correlations for the first-order five-factor model are presented in Table III. Subgroup analyses based on women with confirmed histology only yielded similar results.

Sum scores on each EHP-30 subscale and the total scale were then calculated and correlated with scores on the eight domains of the SF-36 (Table IV). Results suggested that the individual scale scores on the EPH-30 were differentially associated with scores on the individual scales of the SF-36. Moreover, subscales with related content had higher correlations than subscales with unrelated content (e.g. the association between 'pain' and 'bodily pain' was higher than the association between 'emotional well-being' and 'bodily pain').

Internal consistency and test–retest-reliability

Descriptive statistics for the five subscales of EHP-30 are presented in Supplementary Table SII. All corrected item-total correlations were above the recommended value of 0.30 and the average inter-item correlations were between 0.62 and 0.71. Overall, this suggests good internal consistency reliability. EHP-30 also showed acceptable internal consistency for all subscales, with Cronbach's alpha for 'pain' = 0.96, 'control and powerlessness' = 0.94, 'emotional wellbeing' = 0.91, 'social support' = 0.88 and 'self-image' = 0.88.

The agreement between first and second measure of the EHP-30 subscales was examined using Bland–Altman plots (Fig. 2). The plots showed symmetry around the *X*-axis for all subscales. Furthermore, between 45.8% and 73.6% of participants gave the exact same score on pre and post measurement for all EHP-30 items, and between 86.1% and 96.2% of participants scored the same or ±1 value on pre and post-measurement for all EHP-30 items (Supplementary Table SIII).

Convergent validity

Bland–Altman plots representing the agreement between EHP-30 subscales and relevant SF-36 subscales showed no symmetry around the *X*-axis (Fig. 3). Results of the Pearson's correlations in Table IV showed high negative correlations between EHP-30 subscales and related SF-36 subscales.

Table II Fit statistics for the confirmatory factor analysis.

Model	χ^2 (df) *P*	CFI	TLI	RMSEA (90% CI)	SRMR	AIC	BIC
First-order models							
1. One-factor model	2329.955 (405) 0.000	0.731	0.711	0.125 (0.120–0.130)	0.078	22403.404	22737.936
2. Five-factor model	1137.652 (395) 0.000	0.896	0.886	0.079 (0.073–0.084)	0.050	20946.365	21318.068
3. Four-factor model	1539.906 (399) 0.000	0.841	0.826	0.097 (0.092–0.102)	0.069	21430.896	21787.730
Second-order models							
4. Higher-order five-factor model	1163.101 (400) 0.000	0.893	0.884	0.079 (0.074–0.085)	0.054	20965.699	21318.817
5. Higher-order four-factor model	1545.839 (401) 0.000	0.840	0.827	0.097 (0.092–0.102)	0.069	21433.930	21783.330

AIC, Akaike information criterion; BIC, Bayesian information criterion; χ^2, Chi-square; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker–Lewis index.

Table III Factor loadings (standard error) and factor correlations (standard error) for the first-order five-factor model.

	First-order 5-factor model				
EHP-30 items	**PAIN**	**CTRLPW**	**EMO**	**SOC**	**SELF**
Item 1	0.847 (0.019)				
Item 2	0.904 (0.015)				
Item 3	0.813 (0.020)				
Item 4	0.757 (0.025)				
Item 5	0.785 (0.025)				
Item 6	0.860 (0.017)				
Item 7	0.684 (0.032)				
Item 8	0.832 (0.020)				
Item 9	0.907 (0.012)				
Item 10	0.930 (0.010)				
Item 11	0.798 (0.021)				
Item 12		0.809 (0.024)			
Item 13		0.862 (0.024)			
Item 14		0.905 (0.013)			
Item 15		0.816 (0.023)			
Item 16		0.839 (0.024)			
Item 17		0.837 (0.020)			
Item 18			0.848 (0.019)		
Item 19			0.837 (0.022)		
Item 20			0.899 (0.014)		
Item 21			0.821 (0.023)		
Item 22			0.743 (0.031)		
Item 23			0.479 (0.044)		
Item 24				0.807 (0.026)	
Item 25				0.857 (0.026)	
Item 26				0.721 (0.030)	
Item 27				0.835 (0.024)	
Item 28					0.868 (0.023)
Item 29					0.907 (0.021)
Item 30					0.778 (0.038)
PAIN	1.00				
CTRLPOW	0.836 (0.023)	1.00			
EMO	0.720 (0.033)	0.878 (0.020)	1.00		
SOC	0.653 (0.040)	0.791 (0.031)	0.814 (0.031)	1.00	
SELF	0.680 (0.038)	0.735 (0.037)	0.714 (0.042)	0.679 (0.043)	1.00

All factor loadings and factor correlations are statistically significant ($P < 0.001$).
CTRLPW, control and powerlessness; EHP-30, the Endometriosis Health Profile 30 Questionnaire; EMO, emotional wellbeing; SELF, self-image; SOC, social support.

Discussion

The results of the present study support the original five-factor structure of the EHP-30 and thereby the use of five independent scale-scores when applying the questionnaire in clinical and research practice. Lack of agreement between EHP-30 and SF36v2 scale scores demonstrated by Bland–Altman plots indicate that the two questionnaires do not measure the exact same constructs and therefore that the EHP-30 questionnaire measures unique impact of endometriosis on HRQoL not captured by the generic SF-36v2. However, Bland–Altman plots demonstrated agreement between test and retest measures of the EHP-30 subscales suggesting that scoring might be largely unaffected by the cyclical/menstrual pain symptoms in endometriosis. Four out of five participants stated that they continued the same treatment throughout the study period, which corresponds with the stable test–retest result. Overall, results indicate that the EHP-30 is a valid and stable measure of HRQoL in women with endometriosis. Sufficient validity and reliability of the Danish version EHP-30 was also demonstrated.

The study had several strengths. It is the first study to examine several competing factor models of the EHP-30 using CFA and also the

Table IV EHP-30 subscale and total sum correlations with relevant SF-36 domains.

EHP-30 scales[a]	SF-36 Physical functioning	SF-36 Role physical	SF-36 Bodily pain	SF-36 General Health	SF-36 Vitality	SF-36 Social functioning	SF-36 Role emotional	SF-36 Mental health
Pain	−0.710	−0.703	−0.850	−0.561	−0.594	−0.649	−0.483	−0.505
Control and powerlessness	−0.560	−0.642	−0.761	−0.598	−0.706	−0.679	−0.545	−0.651
Emotional wellbeing	−0.397	−0.508	−0.589	−0.519	−0.677	−0.639	−0.649	−0.729
Social support	−0.373	−0.509	−0.541	−0.541	−0.659	−0.571	−0.518	−0.578
Self-image	−0.515	−0.534	−0.542	−0.540	−0.578	−0.545	−0.535	−0.564
Total (sum of all items)	−0.618	−0.681	−0.794	−0.630	−0.724	−0.711	−0.604	−0.672

All correlations are significant at the 0.001 level (two-tailed).
[a]For EHP-30, lower scores indicate better quality of life whereas for the Short Form 36 (SF-36), higher scores indicate better quality of life.

first to apply Bland–Altman plots for evaluation of reliability as well as for convergent validity of this questionnaire. Moreover, it is the first study to validate the EHP-30 in a Danish context. By recruiting patients from multiple centers, a diverse and large sample of endometriosis patients was included in the study. Finally, there was a high response rate. There were also limitations. A larger study, including data from different countries on different continents, would be better designed to exclude potential population bias. This study did not collect data on menstrual cycle stage or hormonal and pain medication. Such data could be used to determine whether scores from the EHP-30 might be unaffected by menstrual cycle stage. In addition, histology could not be confirmed in 33.2% of participants, however, subgroup analysis, based on women with confirmed histology only, yielded similar results.

Item-level analysis of the scores on EHP-30 indicated few missing data, adequate score distribution, and variability and homogeneity within acceptable limits. As missing data are traditionally considered an indicator of items that are difficult to understand, this suggests that the questions were generally understandable and meaningful to women with endometriosis. Overall, the questionnaire appears to generate sufficient score variability to capture differences in levels of QoL among participants. Scores on one item (Item 23) were highly skewed suggesting questionable range variability. This item belongs to the 'Emotional wellbeing' scale. Scrutiny of items on this scale reveals that Item 23 addresses aggressive or violent behavior as a response to endometriosis, whereas the remaining items address negative emotions. The results could suggest that violent or aggressive behavior as a response to endometriosis might be rare or possibly that women are more reluctant to report this type of behavior.

Results from the CFAs indicate that a one-factor model (Model 1) fitted the data poorly. Consequently, items on the EHP-30 do not appear to reflect a single underlying construct as suggested by Jenkinson *et al.* (2008), and scores from the questionnaire are best interpreted as subscale scores. In a similar vein, 'pain' and 'control and powerlessness' items do not appear to reflect the same underlying construct as suggested by Jones *et al.* (2006). Our results showed that two models, a first-order five-factor model and a higher-order five-factor model, were both acceptable in terms of model fit, but they could not be differentiated based on fit values alone. To determine which model represented the data best, difference testing was applied. Results supported the first-order five-factor model, which is in line with the majority of previous studies (American, Chinese, Persian and Dutch) using the more explorative statistical method, PCA (Jones *et al.*, 2001; Jenkinson *et al.*, 2008; Nojomi *et al.*, 2011; van de Burgt *et al.*, 2011; Jia *et al.*, 2013). This finding of a five-factor model across several countries and continents supports the crosscultural validity of the EHP-30. Analyses demonstrated that scores on the individual EHP-30 subscales and scores on the individual SF-36 subscales were differentially correlated, and that subscales with related content were highly correlated (as an example, that the correlation between scores on the SF-36 subscale 'Bodily pain' and the EHP-30 subscale 'Pain' was higher than the correlation between the SF-36 subscale 'Bodily pain' and EHP-30 'total sum score' and also higher than the correlation between the SF-36 subscale 'Bodily pain' and the EHP-30 subscale 'Social support'). The fact that subscales from the two questionnaires with related content had higher correlations than subscales with unrelated content supports the use of scale scores rather than a total sum score when applying the EHP-30. Overall, reasonable fit indices and high factor loadings confirm the original five-factor structure and support the construct validity of this questionnaire.

Post hoc analyses indicated that a residual covariance between Items 22 and 21 would improve the fit of the model. Covariances represent systematic measurement error in item responses and may arise through item characteristics, such as item overlap or the presence of an unmodeled factor (Byrne, 2012). Both items were included in the 'Emotional wellbeing' scale. Further inspection of this scale suggests that whereas several items appear to describe responses to depressed mood, Items 21 and 22 appear to describe responses to mood fluctuations possibly more indicative of cyclothymic or simply just 'everyday' ordinary mood swings. Moving forward, scale developers could consider removing one or more items from the 'emotional wellbeing' scale or rephrase items.

Reliability analyses showed acceptable internal consistency and excellent test–retest reliability for all subscales. To be a reliable and stable measure of QoL in endometriosis, the respondent's answers on the questionnaire must be independent of the monthly fluctuations of menstrual/cyclical pain. Consequently, no control for time of menstrual cycle was included when assessing QoL at test and retest in the present study. Questionnaires were answered randomly within an interval of 6–12 weeks. Therefore, some participants might have had their menstrual bleeding while answering the first questionnaire but not while answering the second questionnaire. Yet, the high agreement between the two measurement points supports the stability and reliability of the EHP-30

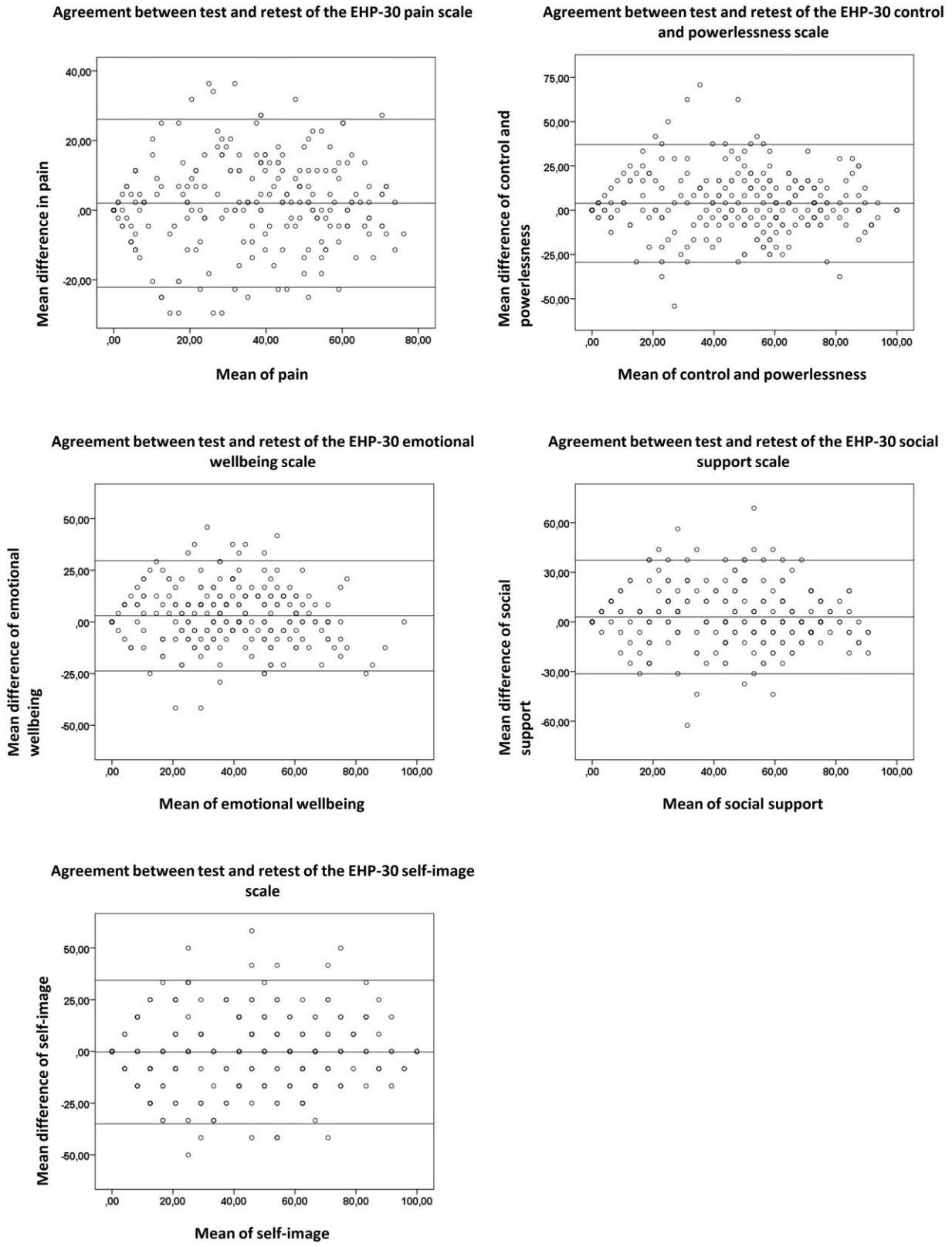

Figure 2. Bland–Altman plots of the agreement between test–retest of the EHP-30 subscales. The Bland–Altman plots are showing the agreement between the measures of the test (the first questionnaire) and the retest (the second questionnaire) of the EHP-30 subscales. Agreement is demonstrated by symmetry in the plot around the 0-axis.

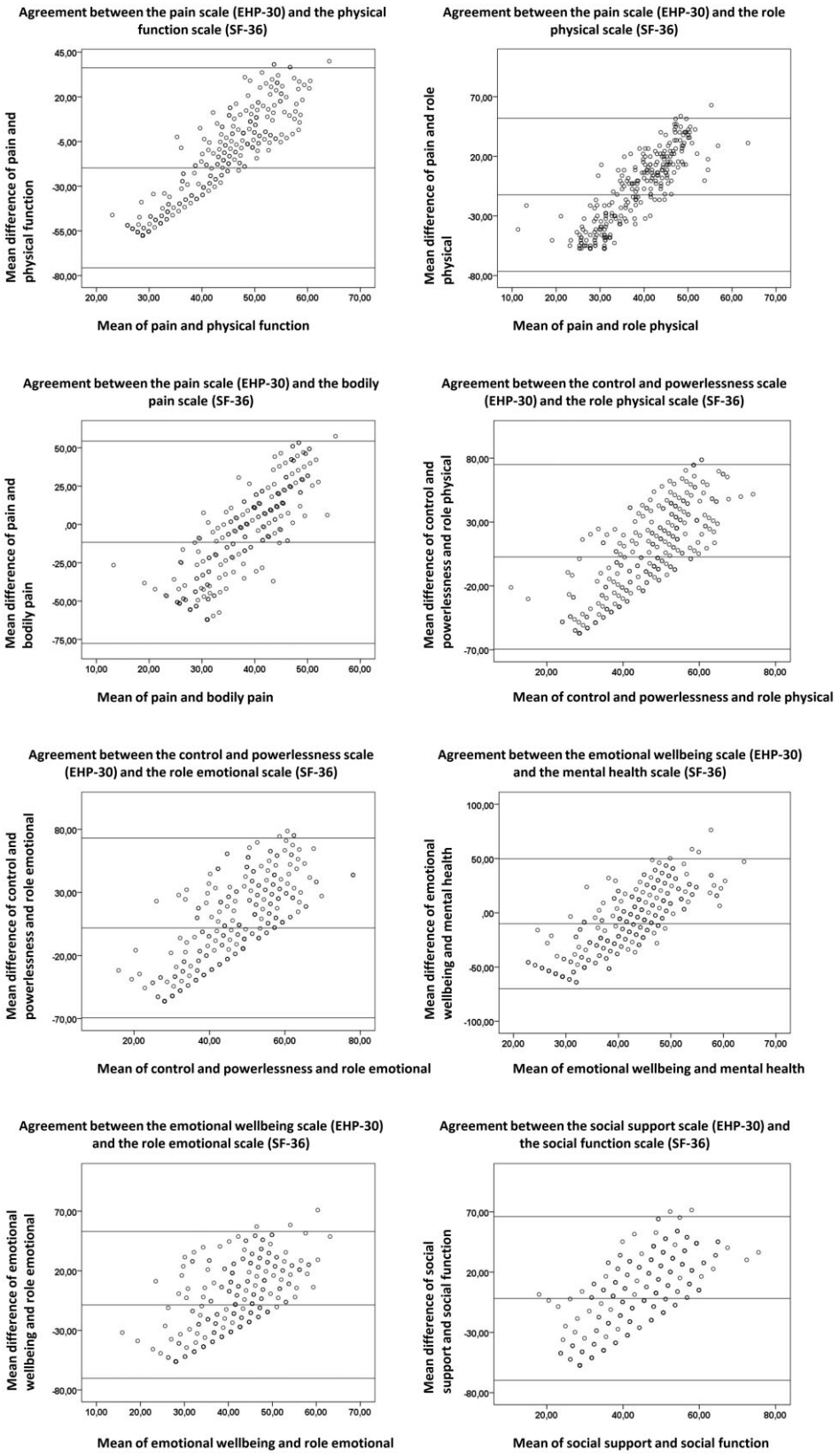

Figure 3. Bland–Altman plots of the agreement between EHP-30 subscales and corresponding SF-36 subscales. The Bland–Altman plots are showing the agreement between the EHP-30 subscales and the corresponding SF-36 subscales. No symmetry in the plot around the 0-axis indicates no agreement between the two measures. SF-36, Short Form 36.

and indicates that scoring might be influenced by endometriosis-related pain in general but appears to be relatively independent of additional pain fluctuations caused by the menstrual cycle. The fact that the vast majority of participants did not change treatment between the two data collection points might also suggest limited impact of cycle stage on questionnaire completion. This is an important result because it indicates that EHP-30 can be used any day during the menstrual cycle without having to control for the time of period when examining endometriosis-related QoL in clinical practice or research studies. Yet, we would advise future researchers to obtain data related to menstrual cycle stage as well as hormonal and pain medication status when completing the EHP-30 for a more thorough assessment of these factors in relation to our result. The visual examination of convergent validity by use of Bland–Altman plots indicated that the EHP-30 subscales do not measure the same constructs as the corresponding SF-36 subscales. These results support the use of a disease-specific QoL instrument to measure the impact of endometriosis on QoL in women with the disease. High correlations between EHP-30 subscales and related SF-36 subscales (e.g. the EHP-30 'pain' scale and the SF-36 'bodily pain' scale) further support the convergent validity of this questionnaire.

EHP-30 can be used to understand and measure the multidimensional impact of endometriosis on QoL in order to improve clinical care pathways. In the Danish public health care system, selected subscales from EHP-30 are being used to monitor and evaluate endometriosis treatment. Once a year, patients are asked to fill out an electronic questionnaire, including relevant parts of EHP-30, that are returned to the endometriosis outpatient clinic. Nurses will then evaluate the responses in order to decide whether patients are in stable and effective treatment, if they need a telephone consultation, or whether they should be seen for a consultation in the outpatient clinic.

Conclusion

The EHP-30 questionnaire measures the unique impact of endometriosis on HRQoL in patients with this disabling disorder and, importantly, scoring might be unaffected by the cyclical/menstrual pain symptoms characteristic of endometriosis. The five-factor structure of the EHP-30, originally suggested by the questionnaire developers, was supported, and thereby the use of five separate scale-scores in research and clinical practice is recommended. The study also demonstrated sufficient validity and reliability of the Danish version EHP-30. Overall, EHP-30 appears to be a valid, stable and specific measure of HRQoL in women with endometriosis, and the finding of a five-factor model across several countries supports the crosscultural validity of the EHP-30.

Supplementary data

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

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Acknowledgements

We thank Vibeke Amelung from Svendborg Hospital for help with recruitment. We also thank The Danish Endometriosis Patients Association for sponsoring the study and for help in recruiting and the Health Research Fund of Central Denmark Region for sponsoring a scholarship.

Authors' roles

K.E.H. coordinated the project. K.E.H., A.F. and U.S.K. conceived and designed the study. K.E.H., K.R., A.G.E. and A.F. executed the study and obtained the data. K.E.H., R.L., H.M. and U.S.K. analyzed and interpreted the data. K.E.H., R.L. and U.S.K. drafted the manuscript and K.R., A.G.E., H.M. and A.F. critically revised the manuscript. All authors read and approved the final submitted manuscript.

Funding

This work was supported by the Danish Endometriosis Association, which is a nongovernmental organization run by women with endometriosis and by a scholarship from the Health Research Fund of Central Denmark Region.

Conflict of interest

The authors have no conflicts of interest.

References

Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000;**25**:3186–3191.

Bjorner JB, Damsgaard MT, Watt T, Groenvold M. Tests of data quality, scaling assumptions, and reliability of the Danish SF-36. *J Clin Epidemiol* 1998a;**51**:1001–1011.

Bjorner JB, Thunedborg K, Kristensen TS, Modvig J, Bech P. The Danish SF-36 Health Survey: translation and preliminary validity studies. *J Clin Epidemiol* 1998b;**51**:991–999.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**327**:307–310.

Bourdel N, Chauvet P, Billone V, Douridas G, Fauconnier A, Gerbaud L, Canis M. Systematic review of quality of life measures in patients with endometriosis. *PLoS One* 2019;**14**:e0208464.

Byrne BM. *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. New York: Routledge, 2012.

Chauvet P, Auclair C, Mourgues C, Canis M, Gerbaud L, Bourdel N. Psychometric properties of the French version of the Endometriosis Health Profile-30, a health-related quality of life instrument. *J Gynecol Obstet Hum Reprod* 2017;**46**:235–242.

Chen LC, Hsu JW, Huang KL, Bai YM, Su TP, Li CT, Yang AC, Chang WH, Chen TJ, Tsai SJ *et al.* Risk of developing major depression and anxiety disorders among women with endometriosis: a longitudinal follow-up study. *J Affective Disord* 2016;**190**:282–285.

Culley L, Law C, Hudson N, Denny E, Mitchell H, Baumgarten M, Raine-Fenning N. The social and psychological impact of endometriosis on women's lives: a critical narrative review. *Hum Reprod Update* 2013;**19**:625–639.

De Graaff AA, D'Hooghe TM, Dunselman GAJ, Dirksen CD, Hummelshoj L, Simoens S, Bokor A, Brandes I, Brodszky V, Canis M, *et al.*; WERF EndoCost Consortium. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. *Hum Reprod* 2013;**28**:2677–2685.

Ferrero S, Arena E, Morando A, Remorgida V. Prevalence of newly diagnosed endometriosis in women attending the general practitioner. *Int J Gynaecol Obstet* 2010;**110**:203–207.

Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ* 1993;**306**:1440–1444.

Gorecki C, Brown JM, Cano S, Lamping DL, Briggs M, Coleman S, Dealey C, McGinnis E, Nelson AE, Stubbs N *et al.* Development and validation of a new patient-reported outcome measure for patients with pressure ulcers: the PU-QOL instrument. *Health Qual Life Outcomes* 2013;**11**:95.

Grundström H, Rauden A, Wikman P, Olovsson M. Psychometric evaluation of the Swedish version of the 30-item endometriosis health profile (EHP-30). *BMC Women's Health* 2020;**20**:204.

Hansen KE, Kesmodel US, Baldursson EB, Kold M, Forman A. Visceral syndrome in endometriosis patients. *Eur J Obstet Gynecol Reprod Biol* 2014;**179**:198–203.

Jackson DL, Gillaspy JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol Methods* 2009;**14**:6–23.

Jenkinson C, Kennedy S, Jones G. Evaluation of the American version of the 30-item Endometriosis Health Profile (EHP-30). *Qual Life Res* 2008;**17**:1147–1152.

Jia SZ, Leng JH, Sun PR, Lang JH. Translation and psychometric evaluation of the simplified Chinese-version Endometriosis Health Profile-30. *Hum Reprod* 2013;**28**:691–697.

Jones G, Jenkinson C, Kennedy S. Evaluating the responsiveness of the Endometriosis Health Profile Questionnaire: the EHP-30. *Qual Life Res* 2004;**13**:705–713.

Jones G, Jenkinson C, Taylor N, Mills A, Kennedy S. Measuring quality of life in women with endometriosis: tests of data quality, score reliability, response rate and scaling assumptions of the Endometriosis Health Profile Questionnaire. *Hum Reprod* 2006;**21**:2686–2693.

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**:258–264.

Marinho MCP, Magalhaes TF, Fernandes LFC, Augusto KL, Brilhante AVM, Bezerra L. Quality of life in women with endometriosis: an integrative review. *J Womens Health (Larchmt)* 2018;**27**:399–408.

Muthén LK, Muthén BO. *Mplus: Statistical Analysis with Latent Variables: User's Guide (Version 8)*. Los Angeles, CA: Authors, 2017.

Nojomi M, Bijari B, Akhbari R, Kashanian M. The assessment of reliability and validity of Persian version of the endometriosis health profile (EHP-30). *Iran J Med Sci* 2011;**36**:84–89.

Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 2001;**66**:507–514.

van de Burgt TJ, Hendriks JC, Kluivers KB. Quality of life in endometriosis: evaluation of the Dutch-version Endometriosis Health Profile-30 (EHP-30). *Fertil Steril* 2011;**95**:1863–1865.

Verket NJ, Andersen MH, Sandvik L, Tanbo TG, Qvigstad E. Lack of cross-cultural validity of the Endometriosis Health Profile-30. *J Endometr Pelvic Pain Disord* 2018;**10**:107–115.

Vigano P, Parazzini F, Somigliana E, Vercellini P. Endometriosis: epidemiology and aetiological factors. *Best Pract Res Clin Obstet Gynaecol* 2004;**18**:177–200.