BMJ Open

Uterine Fibroid Symptom and Quality of Life questionnaire (UFS-QOL NL) in the Dutch population: a validation study

Alieke L Keizer ,[1] Paul J M van Kesteren,[2] Caroline Terwee,[3] Maria E de Lange,[1] Wouter J K Hehenkamp,[1] Helen S Kok[4]

¹Obstetrics and Gynaecology, Amsterdam UMC, Amsterdam, The Netherlands
²Obstetrics and Gynaecology, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands
³Epidemiology and Data Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
⁴Obstetrics and Gynaecology, Alrijne Ziekenhuis Locatie Leiden, Leiden, The Netherlands

Correspondence to
Alieke L Keizer;
i.keizer@amsterdamumc.nl

ABSTRACT

Objective Uterine fibroids can cause a variety of symptoms in women, from heavy menstrual bleeding and dysmenorrhea to bulk symptoms. The Uterine Fibroid Symptom and health-related Quality Of Life questionnaire (UFS-QOL) is a patient-reported outcome measure developed for assessing fibroid-related symptoms in a standardised way. Our aim was to translate and validate the UFS-QOL in Dutch.

Design Validation study.

Setting Patients were recruited by a gynaecologist at the outpatient clinic.

Participants Women with uterine fibroids.

Methods The UFS-QOL was translated into Dutch (UFS-QOL NL) and validated through testing construct validity (comprising of structural validity and hypotheses testing), reliability, responsiveness and interpretability, assessing floor and ceiling effects and minimal important change. An option to answer 'not applicable' was added to the translated questionnaire.

Results 191 women with uterine fibroids completed the UFS-QOL NL at baseline, after 2 weeks and after 3 months. The questionnaire retained the same factor structure after translation (Comparative Fit Index 0.94–0.95; Tucker-Lewis fit Index 0.93–0.95; Root Mean Square Error of Approximation 0.10–0.11) and correlations to other questionnaires (RAND 36, Hospital Anxiety and Depression Scale and Golombok Rust Inventory of Sexual Satisfaction) were generally moderate, as hypothesised (Pearson's r 0.3–0.7). We found a sufficient reliability with intraclass correlation coefficients of approximately 0.8–0.9 for all subscales. Responsiveness was sufficient when testing hypotheses comparing women who had surgery with those who did not. Cronbach's alpha was higher than 0.7 for all subscales, indicating sufficient internal consistency and there were no concerns about floor or ceiling effects. Minimal important change could not be calculated due to low correlation between the different subscales and the anchor question.

Conclusions The results support the measurement properties of the Dutch UFS-QOL for assessing fibroid-related symptoms and health-related quality of life in Dutch women with uterine fibroids.

INTRODUCTION

Uterine fibroids are benign monoclonal tumours arising from a single smooth

Strengths and limitations of this study

► The Uterine Fibroid Symptom and health-related Quality Of Life questionnaire (UFS-QOL) measures fibroid-related symptoms and quality of life in women with uterine fibroids.
► The UFS-QOL is a validated measurement instrument in Dutch in women with uterine fibroids.
► The UFS-QOL is translated and validated in the Dutch population following COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines.
► The UFS-QOL can be used in a research setting and in a clinical setting to evaluate fibroid treatment.

muscle cell. Aetiology is largely unknown but seems to be multifactorial.[1] Incidence is reported to be as high as 77%.[2] Uterine fibroids are symptomatic in up to 50% of women, with symptoms varying from heavy menstrual bleeding to dysmenorrhea, bulk symptoms due to an enlarged uterus and fertility problems.[3 4] Symptoms are generally assessed through anamnesis. For quantifying menstrual blood loss, the Pictorial Blood Assessment Chart (PBAC) can be used.[5] There are no validated questionnaires in Dutch to assess symptom severity in patients with uterine fibroids. The Uterine Fibroid Symptom and Quality Of Life questionnaire (UFS-QOL) has been developed in English and validated in the American population.[6–8] It has been translated in Brazilian Portuguese, Chinese and Spanish.[9–11]

International guidelines for validating questionnaires have been developed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative.[12] These guidelines provide recommendations regarding a systematic translation of a questionnaire and measuring its validity, reliability, responsiveness and interpretability. In short, validity consists of content validity, structural validity

BMJ

Table 1 (A) Construct validity hypotheses; (B) Construct validity results

(A) Construct validity hypotheses			(B) Construct validity results	
Subscale	Correlated item	Suspected correlation*	Correlation between items	Pearson's r
Symptom severity	Self-rated symptom severity	High	Moderate	0.59
Concern	RAND 36 subscale 'general health perception'	Moderate	Moderate	0.31
	HADS subscale 'anxiety'	Moderate	Low negative	−0.21
Activities	RAND 36 subscale 'bodily pain'	High	Moderate	0.52
	RAND 36 subscale 'physical function'	Moderate	Moderate	0.58
Energy/mood	RAND 36 subscale 'vitality'	High	High	0.73
	RAND 36 subscale 'mental health'	Moderate	Moderate	0.60
Control	RAND 36 subscale 'health change'	High	Moderate	0.42
	RAND 36 subscale 'general health perception'	Moderate	Moderate	0.43
Self-conscious	HADS subscale 'depression'	Low	Low negative	−0.06
Sexual functioning	GRISS subscale 'female avoidance'	High negative	Moderate negative	−0.61

*A high correlation was defined as 0.7 or higher. A correlation between 0.7 and 0.3 was defined as moderate. A low correlation was defined as 0.3 or lower.
GRISS, Golombok Rust Inventory of Sexual Satisfaction; HADS, Hospital Anxiety and Depression Scale.

(assessing the dimensionality of the questionnaire) and testing hypotheses regarding internal relationships and relationships with scores from other questionnaires. Reliability and measurement error can be tested through repeating the questionnaire in a timeframe long enough for the patient to not precisely remember which answer they filled out previously, but short enough to prevent symptoms from changing in the meantime. Responsiveness is the ability of a questionnaire to detect change over time in the construct to be measured, for instance after treatment. Interpretability is the extent to which qualitative meaning can be assigned to a questionnaire's quantitative scores or change in scores and can be measured through evaluating the distribution of scores, evaluating possible floor and ceiling effects and estimating a minimal important change (MIC), which is the smallest change in score which patients perceive as important.[13] The aim of this study was to test the validity of the translated UFS-QOL in Dutch women with uterine fibroids.

METHODS
The questionnaire
The UFS-QOL was developed to measure symptoms and health-related quality of life in women with uterine fibroids. The questionnaire contains 37 items, 8 items concerning uterine fibroid symptoms and 29 items concerning health-related quality of life. Patients can answer on a 5-point Likert scale, from 'not at all' to 'a very great deal' and from 'none of the time' to 'all of the time'. Scores are calculated in seven subscales: 'symptom severity' (8 items, 8–40 points), 'concern' (5 items, 5–25 points), 'activities' (7 items, 7–35 points), 'energy/mood' (7 items, 7–35 points), 'control' (5 items, 5–25 points),

'self-conscious' (3 items, 3–15 points), 'sexual function' (2 items, 2–10 points) and can also be summarised into a sum score 'total health-related quality of life' (29 items, 29–145 points). The original version showed sufficient reliability and responsiveness.[7 8] We obtained permission to translate and validate the questionnaire into Dutch from James Spies (Society of the Interventional Radiology).

Translation
The UFS-QOL was translated from English into Dutch by two professional translators and one expert (gynaecologist) with native language Dutch independently. The three translations were combined by the investigators into one final version that all three translators agreed on. This version was translated back into English by two professional translators with native language English independently, to test the linguistic and conceptual equivalence.

Face validity
A committee consisting of the three primary researchers (ALK, PJMvK and HSK) judged the translated questionnaire and deemed it an adequate reflection of symptoms and health-related quality of life in women with uterine fibroids. After consulting with an epidemiologist with expertise in validating questionnaires (CT) an option to fill out 'not applicable' was added to avoid biased results. Items that were scored not applicable were treated as missing. The scoring form was adjusted accordingly (see online supplemental appendices 1; 2 for Dutch questionnaire and scoring form).
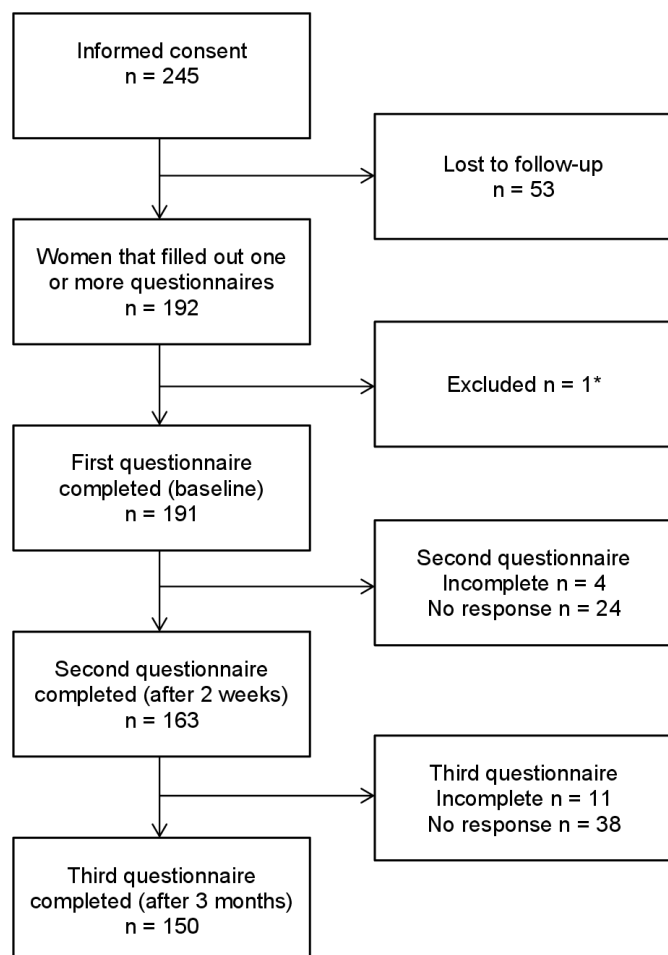
Content validity
An expert panel consisting of three independent gynaecologists and three patients were asked to judge the

translated and adjusted questionnaire regarding relevance and comprehensiveness and comprehensibility, by answering the following questions: Do all items refer to relevant aspects of uterine fibroids? Are all items relevant for the study population? Are all items relevant for measuring symptoms and health-related quality of life in the study population? Is the construct completely covered by all items? A report of the answers and subsequent remarks was drafted and assessed by three researchers (ALK, PJMvK and HSK). Remarks regarding comprehensibility were evaluated and small adjustments in wording were made accordingly, with consent of all three researchers and the professional translators.

Study population

We included women with uterine fibroids, visiting the outpatient clinic of the Alrijne Hospital in Leiden, the Amsterdam UMC location AMC in Amsterdam, the OLVG location Oost hospital in Amsterdam and the Amsterdam UMC, location VUmc, in Amsterdam. Women under the age of 18 years, incapacitated adults and women who did not master the Dutch language were excluded. No further exclusion criteria were formulated.

Figure 1 Flow chart *baseline questionnaire missing.

Patient involvement

Patients were involved as part of the expert panel in testing content validity. They were not involved in other parts of the research.

Study procedures and measures

Women completed a questionnaire, either digitally (online) or on paper. The first questionnaire comprised of several questions about baseline characteristics, the translated UFS-QOL NL, self-rated symptom severity through the question 'I give my fibroid-related symptoms the following grade' (0 meaning no symptoms and 10 meaning the worst symptoms imaginable), the RAND 36-Item Health Survey (RAND 36), measuring physical, mental and social aspects of health-related quality of life (36 items, 8 subscales), the Hospital Anxiety and Depression Scale (HADS), measuring depression and anxiety (14 items, 2 subscales) and the Golombok Rust Inventory of Sexual Satisfaction (GRISS), subscale avoidance, measuring sexual satisfaction, in particular restrictions and avoidance of sexual activities (4 items). The RAND 36, HADS and GRISS were all validated in the Dutch population.[14–16]

Sample size

Sample size requirement was calculated based on factor analysis. For a questionnaire containing 37 items 185 patients should be included (5 patients per item).[12] We aimed to include 190 patients to account for drop out. This sample size would also be sufficient for the other analyses.

Construct validity

Regarding structural validity, a confirmatory factor analysis (CFA) was performed to evaluate the fit of the seven-factor conceptual model. CFA on the polychoric correlation matrix with Weighted Least Squares with Mean and Variance adjustment (WLSMV) estimation was used, using Mplus (V.7). Model fit was evaluated with the Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI) and Tucker-Lewis fit Index (TLI). CFI and TLI ≥0.95 and RMSEA≤0.06 indicated good model fit.[17 18] All other analyses were performed using SPSS (V.23).

Hypotheses were tested regarding construct validity (table 1A). A high correlation was defined as 0.7 or higher. A correlation between 0.7 and 0.3 was defined as moderate. A low correlation was defined as 0.3 or lower. For a sufficient construct validity, 75% of the results should be in accordance with the hypotheses.[19]

Reliability

Internal consistency was calculated using Cronbach's alpha, to assess if multiple items measure the same construct. Cronbach's alpha was considered sufficient if higher than 0.7. Participants completed a second UFS-QOL NL, 2 weeks later at home, either digitally or on paper, to assess test–retest reliability. Reliability was calculated through the intraclass correlation coefficient (ICC,

Table 2 Characteristics of the study population

	Baseline	Two weeks	Three months
Mean age, years	44.5 (SD 6.6, range 25–58) (n=190)	44.8 (SD 6.4, range 25–58) (n=163)	44.9 (SD 6.7, range 25–58) (n=149)
Mean BMI	25.8 (SD 6.0, range 17.3–65.2) (n=191)	25.7 (SD 6.1, range 18.0–65.2) (n=163)	25.8 (SD 6.3, range 18,0–65.2) (n=143)
Race			
African descent	29.8% (n=191)	26.4% (n=163)	26.7% (n=149)
Caucasian	27.2%	28.8%	30.7%
Other European	30.4%	32.5%	33.3%
Asian descent	7.9%	8.6%	6.7%
Mediterranean (Hispanic)	4.7%	3.7%	2.0%
Contraceptive use			
Oral contraceptives (OCS)	37.7% (n=185)	36.8% (n=157)	37.3% (n=144)
Mirena IUD	8.9%	10.4%	10.7%
Copper IUD	2.1%	1.2%	2.0%
OCS +Mirena IUD	10.5%	9.8%	8.7%
OCS +Copper IUD	1.6%	1.8%	2.0%
GnRH analogues	0.5%		
Ulipristal acetate	2.1%	2.5%	2.0%
Depo-Provera	0.5%		
Nuvaring	0.5%		
No hormonal contraceptives or condoms	32.4%	33.7%	33.3%
Missing	3.1%	3.7%	4.0%

BMI, body mass index; GnRH, gonadotropin-releasing hormone; IUD, Intrauterine Device.

two-way mixed effects model for absolute agreement). The limits of agreement (LoA) were calculated as a parameter of measurement error. LoA were calculated as the mean difference between test and retest ±1.96 * SD of the difference. Patients were asked the following extra global question to evaluate whether the patient's health was stable: Has your health status regarding your uterine fibroids changed during the past 2 weeks? Patients who answered yes were excluded from the test–retest analysis. Reliability is considered sufficient if the ICC ≥0.7 and measurement error considered sufficient if LoA were smaller than the MIC.[19]

Responsiveness
Participants completed the UFS-QOL NL after 3 months to test responsiveness. They were asked the following

extra question: Did you have surgery on your uterine fibroids (TransCervical Resection of Myoma (TCRM), myomectomy or hysterectomy) or any other type of intervention during the past 3 months? A distinction was made between women who did not undergo surgery and women who did. The latter group was subdivided into women who underwent a hysterectomy, myomectomy or TCRM. The UFS-QOL consists of seven subscales, for which hypotheses were defined for the different subgroups. Because we consider surgery an effective way of treating heavy menstrual bleeding, we expected that women after surgery would at least have a 50% higher decrease in 'symptom severity' score compared with women who did not have surgery, relative to baseline values. We expected that women who had surgery would at least have a 50%

Table 3 Confirmatory factor analysis

	Root Mean Square Error of Approximation	Comparative Fit Index	Tucker-Lewis fit Index
	Estimate		
Baseline questionnaire	0.104	0.937	0.931
Two-week questionnaire	0.106	0.957	0.953
Three-month questionnaire	0.103	0.950	0.946

Table 4	Internal consistency
Subscale	**Cronbachs α**
Symptom severity	0.79
Concern	0.92
Activities	0.93
Energy/mood	0.94
Control	0.89
Self-conscious	0.74
Sexual functioning	0.89
Total QOL	0.97

QoL, quality of life.

higher increase in 'symptom severity' score compared with women who did not have surgery relative to baseline values in the subgroup 'concern'. We expected that women who had surgery would at least have a 50% higher increase in 'activities' score compared with women who did not have surgery relative to baseline values. We expected that women who had surgery would at least have a 50% higher increase in 'energy/mood' score compared with baseline values compared with women who did not have surgery. We expected that women who had hysterectomy or myomectomy would at least have 50% higher increase in 'self-conscious' score compared with baseline values compared with women who did not have surgery or underwent a TCRM. Finally, we expected that women who had surgery would at least have a 50% higher increase in 'sexual functioning' score compared with baseline values compared with women who did not have surgery. For a sufficient responsiveness, 75% of the results should be in accordance with the hypotheses.[19]

Interpretability

The interpretability of the questionnaire was evaluated by considering the distribution of scores, evaluating floor and ceiling effects and estimating MIC.

Distribution of scores was examined by plotting scores in a histogram.

Floor and ceiling effects were considered present if more than 15% of participants scored the highest or lowest possible score, indicating a possible lack of items at the upper or lower end of the scale.[20]

We used the receiver operating characteristics (ROC) cut-off point to determine an MIC based on the following question in the 3-month questionnaire for women who had surgery (our anchor): How does your current condition regarding your uterine fibroids compare to how it was before surgery? Possible answers were: completely recovered, much improved, moderately improved, slightly improved, unchanged, slightly deteriorated, moderately deteriorated, much deteriorated or worse than ever. Patients who answered completely recovered, much improved or moderately improved were considered as importantly improved. Patients who indicated no change or experienced slight improvement or deterioration were considered not importantly changed. Patients who answered moderately deteriorated, much deteriorated or worse than ever were considered importantly deteriorated. The correlation between the anchor and change in score for women after surgery was calculated using Spearman's rho. This correlation should be at least 0.50 to enable estimation of the MIC. The ROC cut-off point for which the sum of percentages of false-positive classifications (importantly changed according to the questionnaire, but not according to the anchor) and false-negative classifications (not importantly changed according to the questionnaire, but actually so according to the anchor) is smallest was considered the MIC.[21]

RESULTS
Baseline characteristics
In total, 245 women gave informed consent. Eventually 191 women completed the first questionnaire. The

Table 5	Test–retest reliability and measurement error						
				LoA*			
Subscale	**ICC**	**95% CI**		**Mean diff**	**SD diff**	**LoA**	
Symptom severity	0.81	0.70 to 0.87		4.8	12.5	−19.8	29.3
Concern	0.93	0.90 to 0.95		−0.4	12.6	−25.2	24.3
Activities	0.90	0.86 to 0.93		−2.2	12.3	−26.3	22.0
Energy/mood	0.90	0.85 to 0.93		−3.0	12.2	−26.9	20.9
Control	0.84	0.76 to 0.89		−3.9	14.8	−32.9	25.1
Self-conscious	0.76	0.67 to 0.83		−3.7	19.6	−42.2	34.8
Sexual functioning	0.84	0.78 to 0.89		−1.1	18.0	−36.4	34.2
Total QOL score	0.92	0.89 to 0.95		−2.6	9.8	−21.8	16.6

*Range of subscales: 'symptom severity' (8 items, 8–40 points), 'concern' (5 items, 5–25 points), 'activities' (7 items, 7–35 points), 'energy/mood' (7 items, 7–35 points), 'control' (5 items, 5–25 points), 'self-conscious' (3 items, 3–15 points), 'sexual function' (2 items, 2–10 points), 'total health-related quality of life' (29 items, 29–145 points).
ICC, intraclass correlation coefficient; LoA, limits of agreement; QOL, quality of life.

Table 6 Responsiveness

| Subscale | Change in score between baseline and 3-month questionnaires | |
	Surgery Mean (SD)	No surgery Mean (SD)
Symptom severity	−18.7 (26.5)	−10.4 (20.5)
Concern	16.4 (29.0)	4.6 (24.3)
Activities	14.3 (23.9)	2.6 (20.0)
Energy/mood	13.0 (24.9)	4.7 (17.9)
Control	13.0 (26.6)	3.8 (20.8)
Self-conscious*	14.3 (29.1)	3.4 (25.2)
Sexual functioning	10.7 (28.4)	−1.6 (24.5)

*In this subscale, we compared women who underwent abdominal surgery compared with women who underwent hysteroscopic surgery or no surgery.

second questionnaire was filled out by 163 women, 4 women filled it out incompletely and 24 women did not respond. The third questionnaire was filled out by 150 women, 11 women filled it out incompletely, of which 3 were not usable and 38 women did not respond (see also figure 1). For incomplete questionnaires, the scores of different subscales were calculated based on available items where possible. Of the 191 women, 22 women reported a caesarean section and 37 women reported to have undergone other gynaecological surgery previously;

Figure 2 Distribution of sumscores. QOL, quality of life.

20 myomectomy (laparoscopic and hysteroscopic), 1 embolisation, 1 EUG, 4 curettage, 3 colposcopy, 2 polypectomy and 2 (unsuccessful) endometrial ablation. For other baseline characteristics, see table 2.

Construct validity

The original factor structure seemed to be preserved after translation (table 3). CFI and TLI values were close to the criterion of 0.95, RMSEA values were slightly higher than the criterion of 0.06. In total, 55% of the analyses matched the predefined hypotheses (table 1B).

Reliability

Cronbach's alpha was >0.7 for all subscales (table 4). The questionnaire showed sufficient reliability (all ICC's>0.7) for all subscales (table 5). The LoA was −21.8 to 16.6 for the total score.

Responsiveness

Women who had surgery had a higher change in scores when comparing the baseline and 3-month questionnaires than women who did not have surgery, as predicted. Women who underwent surgery had a larger decrease of symptoms than women who did not have surgery, however, the difference was less than the 50% we predicted. On the other QOL-related subscales, women who underwent surgery had a larger increase in QOL than women who did not have surgery. For results, see table 6. In total, 86% of the analyses matched the predefined hypotheses.

Interpretability

We found a normal distribution of sumscores for all questionnaires (figure 2). The percentage of patients with minimum or maximum score on one of the subscales is presented in table 7, for the baseline, 2 weeks and 3 months questionnaire. These result indicate a possible ceiling effect for concern after 3 months and a ceiling effect for sexual functioning.

After 3 months, 57 out of 153 women that filled out the questionnaire reported to have had surgery for uterine fibroids, 27 women underwent hysterectomy, 20 women underwent laparoscopic myomectomy, 10 women underwent hysteroscopic myomectomy. According to the anchor question, one of the women who had surgery deteriorated. The correlations between the anchor and change in score for women after surgery were lower than 0.50 for all subscales (table 8). Therefore, MIC could not be estimated.

DISCUSSION
Main finding

This study supports the validity of the UFS-QOL in Dutch in women with uterine fibroids. The results of this study are comparable to the validation of the original questionnaire and to the translation and validation in other languages.
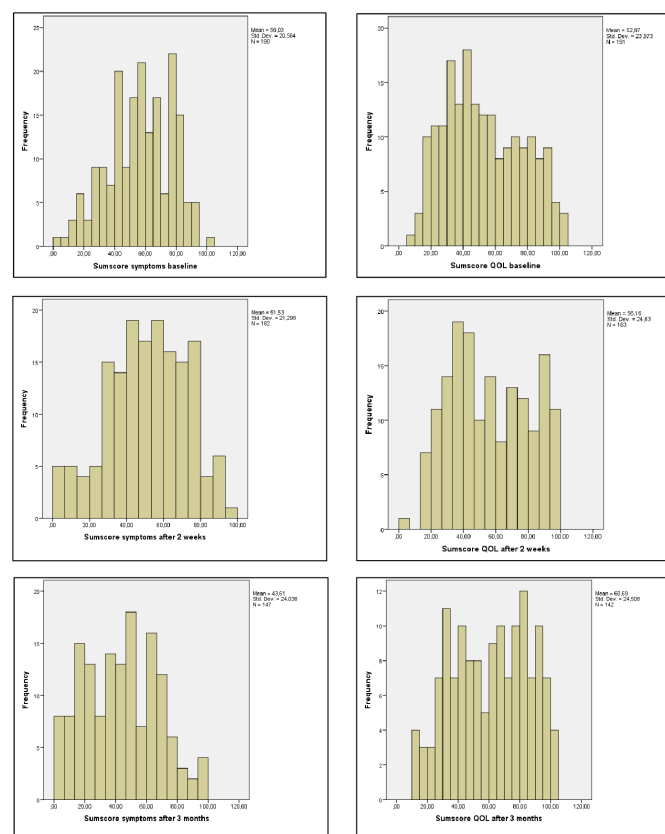
Table 7 Percentage of patients with minimum or maximum score (floor and ceiling effects)

Subscale	Total N	N (%) min. score	N (%) max. score
Baseline			
Symptom severity	190	1 (0.5)	1 (0.5)
Concern	187	7 (3.7)	18 (9.6)
Activities	190	1 (0.5)	14 (7.4)
Energy/mood	189	3 (1.6)	14 (7.4)
Control	191	2 (1.0)	8 (4.2)
Self-concious	189	8 (4.2)	11 (5.8)
Sexual functioning	177	9 (5.1)	33 (18.6)
Total QOL	191	0 (0.0)	3 (1.6)
Two weeks			
Symptom severity	162	0 (0.0)	0 (0.0)
Concern	160	9 (5.6)	21 (13.1)
Activities	163	1 (0.6)	18 (11.0)
Energy/mood	162	1 (0.6)	11 (6.8)
Control	162	1 (0.6)	7 (4.3)
Self-concious	160	5 (3.1)	15 (9.4)
Sexual functioning	151	14 (9.3)	26 (17.2)
Total QOL	163	1 (0.6)	3 (1.8)
Three months			
Symptom severity	147	4 (2.7)	2 (1.4)
Concern	138	3 (2.2)	25 (18.1)
Activities	142	1 (0.7)	16 (11.3)
Energy/mood	142	0 (0.0)	16 (11.3)
Control	142	1 (0.7)	13 (9.2)
Self-concious	144	5 (3.5)	19 (13.2)
Sexual functioning	129	10 (7.8)	23 (17.8)
Total QOL	142	0 (0.0)	4 (2.8)

QOL, quality of life.

Table 8 Correlation between anchor* and change in score after 3 months (after surgery)

Subscale	N	Spearman's Rho
Symptom severity†	53	0.03
Concern‡	49	−0.04
Activities‡	51	−0.09
Energy/mood‡	51	−0.12
Control‡	51	−0.18
Self-conscious‡	50	−0.16
Sexual functioning‡	48	−0.04
Total QOL‡	51	−0.14

*A low score meaning much better, a high score meaning much worse.
†A low score meaning less symptoms, a high score meaning more symptoms.
‡A low score meaning low QOL, a high score meaning high QOL.
QOL, quality of life.

Strengths and limitations

Regarding structural validity, we found CFI and TLI values higher than the minimum criteria of 0.95. However, the RMSEA was higher than the maximum criterion of <0.06. This is often found for patient-reported outcome measures. The RMSEA statistic is found to be problematic for assessing dimensionality of health concepts.[22]

Construct validity was evaluated by comparing the different subscales of the questionnaire with already validated questionnaires in Dutch (RAND 36, HADS, GRISS). Only 55% the results were in accordance with hypotheses. This was also found during the validation of the original questionnaire and translation and validation of the questionnaire in the Brazilian Portuguese and Chinese language.[6 9 10] Possibly the questionnaires we have chosen to test construct validity are to generic and do not capture the symptoms that women with uterine fibroids experience, or the hypotheses we formed about

these questionnaires could have accounted for this more accurately.

Cronbach's alpha was higher than 0.9 on a number of scales, indicating a possible redundancy in items.[19] The questionnaire showed sufficient reliability (all ICC's≥0.7) on all subscales.

In assessing responsiveness, we evaluated the difference between operated and non-operated women. However, we did not take medication use into account. Also we did not investigate the association between clinical variables, like fibroid type, size or number or PBAC score, to the UFS-QOL symptom score. PBAC score was evaluated in the Chinese translation and validation of the questionnaire and showed a positive correlation with UFS-QOL symptom scores.[10] Possibly, it would have provided a larger difference in change of score between operated and non-operated women if we would also have taken into account any medication that was started after the initial questionnaire for all women because these could have also decreased symptoms, or if we would have corrected for fibroid parameters like type, size or vascularity.

Interpretation

We translated and validated an existing questionnaire, rather than developing a new one to measure fibroid-related complaints, so that results of future research in the Dutch population can be compared with the results of international research. An option to answer not applicable on a questionnaire should always be considered and could possibly be added to the original questionnaire.

MIC could not be calculated because the correlations between the anchor and change in score were too low. MIC should be assessed in a future study in a larger group of patients with a greater variety of treatments, and with using multiple anchors.

CONCLUSIONS

The UFS-QOL NL is a standardised measurement instrument for assessing uterine fibroid-related symptoms and health-related quality of life in the Dutch population, with sufficient structural validity, reliability and responsiveness. The UFS-QOL NL can be used to assess symptom severity in prospective research as well as in a clinical setting, as a standardised tool to evaluate subjective symptoms.

ORCID iD
Alieke L Keizer http://orcid.org/0000-0003-2612-563X

REFERENCES

1. Leppert PC, Al-Hendy A, Baird DD, *et al*. Summary of the proceedings of the basic science of uterine fibroids meeting: new developments February 28, 2020. *F S Sci* 2021;2:88–100.
2. Stewart EA, Cookson CL, Gandolfo RA, *et al*. Epidemiology of uterine fibroids: a systematic review. *BJOG* 2017;124:1501–12.
3. Stovall DW. Clinical symptomatology of uterine leiomyomas. *Clin Obstet Gynecol* 2001;44:364–71.
4. Gupta S, Jose J, Manyonda I. Clinical presentation of fibroids. *Best Pract Res Clin Obstet Gynaecol* 2008;22:615–26.
5. Janssen CA, Scholten PC, Heintz AP. A simple visual assessment technique to discriminate between menorrhagia and normal menstrual blood loss. *Obstet Gynecol* 1995;85:977–82.
6. Spies JB, Coyne K, Guaou Guaou N, *et al*. The UFS-QOL, a new disease-specific symptom and health-related quality of life questionnaire for leiomyomata. *Obstet Gynecol* 2002;99:290–300.
7. Harding G, Coyne KS, Thompson CL, *et al*. The responsiveness of the uterine fibroid symptom and health-related quality of life questionnaire (UFS-QOL). *Health Qual Life Outcomes* 2008;6:99.
8. Coyne KS, Margolis MK, Bradley LD, *et al*. Further validation of the uterine fibroid symptom and quality-of-life questionnaire. *Value Health* 2012;15:135–42.
9. Oliveira Brito LG, Malzone-Lott DA, Sandoval Fagundes MF, *et al*. Translation and validation of the uterine fibroid symptom and quality of life (UFS-QOL) questionnaire for the Brazilian Portuguese language. *Sao Paulo Med J* 2017;135:107–15.
10. Yeung SY, Kwok JWK, Law SM, *et al*. Uterine fibroid symptom and health-related quality of life questionnaire: a Chinese translation and validation study. *Hong Kong Med J* 2019;25:453–9.
11. Calaf J, Palacios S, Cristóbal I, *et al*. Validation of the Spanish version of the uterine fibroid symptom and quality of life (UFS-QoL) questionnaire in women with uterine myomatosis. *Med Clin* 2020;154:207–13.
12. Mokkink LB, Terwee CB, Patrick DL, *et al*. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49.
13. de Vet HCW, Terwee CB, Mokkink LB, *et al*. *Measurement in medicine: a practical guide*. Cambridge University Press, 2011.
14. VanderZee KI, Sanderman R, Heyink JW, *et al*. Psychometric qualities of the Rand 36-Item health survey 1.0: a multidimensional measure of general health status. *Int J Behav Med* 1996;3:104–22.
15. Spinhoven P, Ormel J, Sloekers PP, *et al*. A validation study of the hospital anxiety and depression scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997;27:363–70.
16. ter Kuile MM, van Lankveld JJ, Kalkhoven P, *et al*. The Golombok rust inventory of sexual satisfaction (GRISS): psychometric properties within a Dutch population. *J Sex Marital Ther* 1999;25:59–71.
17. Hu Li-tze, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1999;6:1–55.
18. Schreiber JB, Nora A, Stage FK, *et al*. Reporting structural equation modeling and confirmatory factor analysis results: a review. *J Educ Res* 2006;99:323–38.
19. Prinsen CAC, Mokkink LB, Bouter LM, *et al*. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57.
20. Terwee CB, Bot SDM, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
21. de Vet HCW, Ostelo RWJG, Terwee CB, *et al*. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2007;16:131–42.
22. Steven PR, Richard S, Keith FW. Multidimensionality and structural coefficient bias in structural equation modeling: a Bifactor perspective. *Educational and Psychological Measurement* 2012;73:5–26.