VALUE IN HEALTH 15 (2012) 135–142

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Further Validation of the Uterine Fibroid Symptom and Quality-of-Life Questionnaire

Karin S. Coyne, PhD, MPH[1,], Mary Kay Margolis, MPH, MHA[1], Linda D. Bradley, MD[2], Richard Guido, MD[3], G. Larry Maxwell, MD, MC[4], James B. Spies, MD, MPH[5]*

[1]*United BioSource Corporation, Bethesda, MD, USA;* [2]*Cleveland Clinic Department of Obstetrics and Gynecology, Cleveland, OH, USA;* [3]*Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh, Pittsburgh, PA, USA;* [4]*Walter Reed Army Medical Center, Department of Obstetrics and Gynecology, Washington, DC, USA;* [5]*Department of Radiology, Georgetown University Hospital, Washington, DC, USA*

ABSTRACT

Objective: To further examine the reliability, validity and responsiveness of the uterine fibroid symptom and quality-of-life (UFS-QOL) questionnaire among women with and without uterine fibroids. **Methods:** A multicenter, non-randomized, prospective study was conducted with women undergoing treatment for uterine fibroids (fibroid treatment group [FTG]) and normal controls (normal control group [NCG]). Women in the FTG were recruited when they were scheduled for treatment; women in the NCG were recruited during their annual exam. Participants completed the UFS-QOL and a short form 36 health survey (SF-36) at enrollment and at 6 and 12 months. Descriptive statistics, Cronbach's alpha, Spearman's correlations, *t* tests, and general linear models were used to analyze the internal consistency and test–retest reliability, concurrent and discriminant validity, and responsiveness of the UFS-QOL. **Results:** There were 89 NCG and 234 FTG women who completed the study. Mean age was 43.1 years for FTG and 40.8 for NCG ($P < 0.001$). The FTG reported significantly greater symptom severity and worse health-related quality of life (HRQL) than the NCG (all UFS-QOL subscales $P < 0.001$). The UFS-QOL subscales were significantly correlated in the expected direction and magnitude with each SF-36 subscale in the FTG, indicating acceptable concurrent validity. Cronbach's alphas were 0.73 to 0.97, reflecting adequate internal consistency. Each UFS-QOL subscale was responsive to changes after treatment in the FTG with effect sizes ranging between 1.1 and −2.35. The UFS-QOL remained stable in the NCG during the 1 year follow-up. **Conclusion:** The UFS-QOL is a valid and reliable measure to assess symptoms and HRQL in women with uterine fibroids and is highly responsive to treatment-related changes.

Keywords: health-related quality of life, Patient-reported outcomes, Symptoms, Uterine fibroids.

Introduction

Uterine leiomyomata, or fibroids, are among the most common benign tumors of the female reproductive tract, with fibroids identified in up to 77% of hysterectomy specimens [1]. Nearly half of women with uterine fibroids have heavy bleeding [2,3]; many also report moderate to severe abdominal pain [4], urinary frequency and urgency, low back pain, and pain during intercourse [2]. Women with severe bleeding may stay close to a restroom on menstruating days leading to disruption in work productivity, sleep, and social activities [2]. As symptoms and their impact on health-related quality of life (HRQL) and activities of daily living are the primary indications for uterine fibroid therapy, patient-reported outcome measures are the most appropriate tools to measure the impact and outcomes of interventions.

The uterine fibroid symptom and quality-of-life questionnaire (UFS-QOL) is a fibroid-specific symptom and quality of life questionnaire, developed to assess outcomes from fibroid therapies [5]. The initial validation demonstrated its ability to discriminate women without fibroids from women with fibroids and from patients with varying degrees of fibroid symptoms. However, the original validation was cross-sectional and responsiveness analyses could not be conducted. Although one study assessed the responsiveness of the UFS-QOL using a series of patients treated with magnetic resonance–-guided focused ultrasound surgery [6], this study did not have normal controls. Because the UFS-QOL has been used in several fibroid studies [7–10], the purpose of this study was to further evaluate the reliability, validity, and responsiveness of the UFS-QOL among women with and without uterine fibroids.

Materials and Methods

This was a multicenter comparative, prospective outcome study of women undergoing treatment for uterine fibroids, with a normal control group for which the clinical findings have been previously described [11]. Patients undergoing fibroid treatment were enrolled at three university medical centers and one academic military medical center; women without fibroids (normal controls) were recruited at three of these centers along with two private

* *Address correspondence to:* Karin S. Coyne, PhD, MPH, United BioSource Corporation, Outcomes Research, 7101 Wisconsin Avenue, Suite 600, Bethesda, MD 20814.

E-mail: karin.coyne@unitedbiosource.com.

1098-3015/$36.00 – see front matter

doi:10.1016/j.jval.2011.07.007

practices in the Washington, DC metropolitan area. Each center's Institutional Review Board approved the study. The study was conducted in compliance with the Health Information Portability and Accountability Act and the principles outlined in the Declaration of Helsinki. All patients provided written informed consent for study participation.

Recruitment

Women with uterine fibroids were recruited to participate in this study after their treatment procedure was scheduled. Each participant in the fibroid treatment group (FTG) had a routine history and physical examination, supplemented with imaging as appropriate, to confirm her diagnosis. Women in the normal control group (NCG) were recruited as they completed routine well-woman evaluations by their gynecologists. Pre-menopausal women ≥ 30 and ≤ 50 years old who were willing to provide written informed consent and able to speak and read English were included in both groups. Additionally, the women in the FTG had to be scheduled to undergo hysterectomy, myomectomy, or uterine embolization. Women in the NCG had to have no history of uterine fibroids and a normal gynecologic examination with regular menstrual cycles at the time of enrollment. Women were excluded if they were: currently pregnant, had cognitive or psychiatric impairment that would interfere with completing the questionnaires, had comorbidity with life expectancy less than 1 year, or were an active military duty member.

Questionnaire assessment

All participants completed the UFS-QOL, the medical outcomes study short form 36 (SF-36), and a brief sociodemographic questionnaire prior to their scheduled treatment (for FTG) or at enrollment (for NCG). The participants mailed the completed questionnaire packet in a pre-addressed stamped envelope to the data center.

Follow-up questionnaire packets were mailed to each participant at 6 and 12 months. The initial study design was to include a 5-year follow-up period; however, due to a limitation of funding, the study was terminated after 1 year of follow-up. The following questionnaires were completed at both follow-ups: overall treatment effect (OTE) scale, UFS-QOL (or UFS-QOL-hysterectomy), SF-36, and a status form, which collected healthcare utilization and pregnancy status. Each questionnaire is described in greater detail below. Participants mailed the completed questionnaires in pre-addressed stamped envelopes to the data center. Each participant was mailed $20 upon receipt of each questionnaire packet.

Questionnaires utilized in study

UFS-QOL

The UFS-QOL was developed from focus groups of women with leiomyomata. The final questionnaire consists of an 8-item symptom severity scale and 29 HRQL questions, which comprise the following six subscales: concern, activities, energy/mood, control, self-consciousness, and sexual function [5]. Participants are instructed to consider their experiences with uterine fibroids during the previous 3 months. Response options for the symptom severity subscale range from "not at all" to "a very great deal." Response options for the HRQL subscales range from "none of the time" to "all of the time." The Symptom Severity subscale and the HRQL subscale scores of the UFS-QOL are opposite, with higher symptom scores indicating greater symptom severity while higher HRQL subscale scores indicate better HRQL.

UFS-QOL-hysterectomy

Because the wording of the UFS-QOL is based on the presence of uterine fibroids, this questionnaire was modified to allow women with hysterectomies to respond. The UFS-QOL-hysterectomy is the same as the UFS-QOL questionnaire except that the introductory paragraph was changed to instruct the participant to consider each symptom as it relates to how she feels now "*after your hysterectomy*" as opposed to related to her symptoms of her *uterine fibroids or menstrual cycle during the previous 3 months*. The lead-in to each question was changed to: "*Since your hysterectomy. . .*" This is the first time this questionnaire has been implemented.

Medical outcomes SF-36

The medical outcomes SF-36 is a 36-item self-administered generic measure designed to assess general health status [12]. The SF-36 is comprised of the following eight subscales: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health; in addition, it also has two composite scores, the physical component score and mental component score. Individual items from each subscale are combined to form a subscale rating and transformed to a 0 to 100 scale. Higher scores indicate better quality of life and the recall period is 4 weeks. Reference values derived in a healthy population, distributed by age and gender, are available [12].

OTE scale

The OTE is an assessment of change in symptoms since the previous questionnaire completion and was completed at each follow-up visit. The participant indicated on the first question whether her uterine fibroid symptoms (for the FTG) or menstrual symptoms (for the NCG) have improved, remained the same, or worsened since the last visit. If her symptoms improved, she rated the degree of improvement on a 7-point scale from 1 "almost the same, hardly better at all" to 7 "a very great deal better." If her symptoms worsened, she rated the degree of deterioration on a 7-point scale from -1 "almost the same, hardly worse at all" to -7 "a very great deal worse." The OTE was utilized to identify stable patients for the test–retest reliability analysis and to assess responsiveness of the UFS-QOL. Participants scoring -1, 0, or 1 were considered "unchanged"; those scoring less than -1 were considered "worse," and those scoring greater than 1 were considered "improved."

Test–retest reliability substudy

To further evaluate the reliability of the UFS-QOL, a test–retest reliability substudy was conducted. A block randomization scheme of 1:4 for the FTG and 1:3 for the NCG was developed prior to study initiation to randomize 100 FTG women and 30 NCG women to participate in a test–retest reliability substudy. Participants were randomized upon return of their completed 6-month follow-up questionnaires. Participants randomized to the retest visit completed the retest visit approximately 2 weeks (14 days $\pm$ 2 days) after the 6-month follow-up visit. The retest questionnaire packet was the same as the 6-month packet and were completed and returned in the same manner as the others.

Sample size calculation and statistical analysis

An initial recruitment rate of 100 patients per group was planned, based on the within-individual change over time for the treatment groups and reported elsewhere [11]. Scoring of the questionnaires was performed according to the developers' guidelines. Analyses were performed on SAS version 9.1.3 (SAS Institute Inc., Cary, NC); all statistical tests were identified *a priori* and no data imputations were performed for missing data. All statistical tests were two-tailed and were conducted with Type I error probability fixed at 0.05.

Descriptive analyses were performed on patient sociodemographic and clinical characteristics. Additional psychometric analyses were conducted on the UFS-QOL to examine internal consistency reliability, and discriminant and concurrent validity. A Cronbach's alpha ≥ 0.70 was considered to demonstrate internal consistency [13]. Intraclass correlation coefficients (ICC) and Spearman's correlations were calculated to evaluate the degree of association between mean UFS-QOL scores at the 6-month follow-up and 2-week retest visit with scores of 0.70 considered adequate test–retest reliability [14]. Spearman correlations were used to evaluate convergent validity with a correlation range of 0.40 to 0.70 considered adequate.

Paired *t*-tests were used to evaluate whether there was statistically significant change in scores on the UFS-QOL between the test–retest evaluations. Questionnaire change scores, effect size (ES), and standard error of measurement (SEM) from baseline to the 6-month visit and from baseline to the 1-year visit were calculated by treatment group to determine the responsiveness of the UFS-QOL. ES was interpreted as small (0.20), moderate (0.50), or large (0.80) following the guidelines proposed by Cohen [15] in 1988. The SEM is computed by multiplying the standard deviation of the measure by the square root of one minus its reliability coefficient. Scheffe's post-hoc comparisons were performed for general linear models with multiple statistical comparisons. Confirmatory factor analyses (CFA) were performed on the previously determined UFS-QOL subscales using EQS 6.1 (Build 90) for Windows (Multivariate Software Inc., Encino, CA). Bentler's confirmatory fit index (CFI) of ≥ 0.90 was used as the criteria to assess goodness of model fit [16].

Results

Participants

A total of 101 women in the NCG and 274 women in the FTG completed the baseline packets. In the NCG, 89 (88%) subsequently completed the 12-month follow-up visit. In the FTG, 274 com-

Table 1 – Baseline participant demographic characteristics.

Characteristic	NCG (N = 101)	FTG (N = 274)	*P* value
Age (mean, SD years)	40.8 (4.9)	43.1 (4.5)	<0.0001
Race (n, % yes)			
White	43 (42.6%)	120 (43.8%)	0.6159
Black	49 (48.5%)	122 (44.5%)	
Asian	1 (1.0%)	7 (2.6%)	
Hispanic	6 (5.9%)	11 (4.0%)	
Other*	1 (1.0%)	10 (3.6%)	
Missing	1 (1.0%)	4 (1.5%)	
Employment status (n, % yes)			
Employed, full time	65 (64.4%)	217 (79.2%)	0.0061
Employed, part time	18 (17.8%)	20 (7.3%)	
Homemaker	13 (12.9%)	18 (6.6%)	
Student	1 (1.0%)	1 (0.4%)	
Unemployed	2 (2.0%)	9 (3.3%)	
Disabled	1 (1.0%)	9 (3.3%)	
Highest level of education (n, % yes)			
Elementary/primary school	0 (0.0%)	1 (0.4%)	0.3118
Secondary/high school	7 (6.9%)	36 (13.1%)	
Some college	26 (25.7%)	79 (28.8%)	
College degree	35 (34.7%)	92 (33.6%)	
Postgraduate degree	32 (31.7%)	66 (24.1%)	
Comorbid medical conditions (n, % yes)			
None	70 (69.3%)	133 (48.5%)	0.0003
Arthritis	6 (5.9%)	23 (8.4%)	0.4301
Asthma	9 (8.9%)	24 (8.8%)	0.9633
Cancer	0 (0.0%)	2 (0.7%)	0.3893
COPD/emphysema	0 (0.0%)	3 (1.1%)	0.2911
Diabetes	1 (1.0%)	6 (2.2%)	0.4464
Heart problems	0 (0.0%)	3 (1.1%)	0.2911
Hypertension	7 (6.9%)	47 (17.2%)	0.0124
Other†	13 (12.9%)	64 (23.4%)	0.0257
Height (mean, SD feet)	5.4 (0.2)(n = 100)	5.4 (0.2)	0.1062
Weight (mean, SD pounds)	163.7 (37.7)(n = 100)	168.9 (44.5)	0.3043
Body mass index (mean, SD)	27.1 (6.1)(n = 100)	28.2 (6.9)(n = 273)	0.1529

COPD, chronic obstructive pulmonary disease; FTG, fibroid treatment group; NCG, normal control group; SD, standard deviation.

* Other includes American Indian and mixed race.

† Other includes Addison's disease, allergies, anemia, anorexia, anxiety, back problems, breast cancer, bulimia, chronic fatigue syndrome, cervical dysplasia, Crohn's disease, epilepsy, eczema, gastroesophageal reflux disease, glaucoma, hepatitis C, high cholesterol, hypertension, hyperthyroidism, hypothyroidism, lupus, mitral valve prolapse, major depression, Meniere's disease, migraines, muscle tension, obsessive compulsive disorder, osteoporosis, polycystic kidney disease, pulmonary hyalinizing granulomas, recent foot and hand surgery, rheumatoid arthritis, sleep apnea, and spinal/neck degeneration.

pleted the baseline packet; of these, 107 had embolization treatment, 61 had a myomectomy, and 106 had a hysterectomy. There were 89 (83%), 55 (90%), and 91 (86%) women who completed the 12-month follow-up visit within the embolization, myomectomy, and hysterectomy groups, respectively.

The mean age of women in the FTG was significantly higher than the women in the NCG (43.1 versus 40.8 years, $P < 0.001$) (Table 1). There were no significant differences between the groups in terms of race or education. Women in the FTG were slightly more likely to be employed full-time ($P = 0.0061$), whereas women in the NCG were more likely to report no comorbid conditions (69.3% vs. 48.5%, $P = 0.0003$) than the women in the FTG.

Among the NCG, the UFS-QOL item means ranged from 1.1 to 2.1, indicating a very low level of symptoms and HRQL impact (Table 2). Within the FTG, the item means reflected greater severity and impact, ranging from 2.6 to 4.1 (Table 2). At baseline, all individual item means of the UFS-QOL were significantly different between the NCG and FTG (Table 2; all $P < 0.001$). The difference in means was largest (mean difference = 2.5) for item 15 (*made you concerned about soiling underclothes*) and item 28 (*made you feel inconvenienced about always carrying extra pads, tampons, and clothing to avoid accidents*). The smallest mean difference (1.4) was observed for item 24 (*made you feel downhearted and blue*).

At baseline, all subscales of the UFS-QOL and SF-36 differed significantly between the FTG and the NCG (all $P < 0.001$) (Table 3), with the FTG reporting significantly greater symptom severity and lower HRQL on the UFS-QOL and SF-36. Comparisons among the three FTGs revealed no significant differences at baseline, with the exception of the mental health subscale of the SF-36 – the hysterectomy group had a significantly worse mental health score when compared with the UFE group (58.3 vs. 67.8, respectively; $P < 0.01$).

Cronbach's alphas were calculated for each UFS-QOL subscale for the NCG, FTG, and also for each of the three FTGs. The alphas ranged from 0.73 to 0.97 at baseline reflecting strong internal con-

Table 2 – Baseline comparison of UFS-QOL items for NCG and FTG.

UFS-QOL item	Mean (SD)		P value
	NCG (N = 98)	FTG (N = 274)	
1. Heavy bleeding during your menstrual period	1.8 (1.0)	4.1 (1.2)	<0.0001
2. Passing blood clots during your menstrual period	1.6 (0.9)	3.7 (1.3)	<0.0001
3. Fluctuation in the duration of your menstrual period compared to your previous cycle	1.3 (0.6)	3.3 (1.4)	<0.0001
4. Fluctuation in the length of your monthly cycle compared to your previous cycles	1.4 (0.7)	3.3 (1.4)	<0.0001
5. Feeling tightness or pressure in your pelvic area	1.6 (0.9)	3.7 (1.2)	<0.0001
6. Frequent urination during the daytime hours	1.6 (0.9)	3.5 (1.3)	<0.0001
7. Frequent nighttime urination	1.5 (0.9)	3.2 (1.4)	<0.0001
8. Feeling fatigued	2.1 (1.0)	3.9 (1.1)	<0.0001
9. Made you feel anxious about the unpredictable onset or duration of your periods?	1.2 (0.6)	3.4 (1.3)	<0.0001
10. Made you anxious about traveling?	1.2 (0.6)	3.2 (1.4)	<0.0001
11. Interfered with your physical activities?	1.2 (0.6)	3.5 (1.2)	<0.0001
12. Caused you to feel tired or worn out?	1.5 (0.9)	3.8 (1.1)	<0.0001
13. Made you decrease the amount of time you spent on exercise or other physical activities?	1.3 (0.7)	3.6 (1.2)	<0.0001
14. Made you feel as if you are not in control of your life?	1.1 (0.4)	3.3 (1.4)	<0.0001
15. Made you concerned about soiling underclothes?	1.5 (1.0)	4.0 (1.2)	<0.0001
16. Made you feel less productive?	1.3 (0.6)	3.4 (1.2)	<0.0001
17. Caused you to feel drowsy or sleepy during the day?	1.4 (0.8)	3.3 (1.3)	<0.0001
18. Made you feel self-conscious of weight gain?	1.3 (0.7)	3.5 (1.4)	<0.0001
19. Made you feel that it was difficult to carry out your usual activities?	1.2 (0.6)	3.3 (1.3)	<0.0001
20. Interfered with your social activities?	1.2 (0.4)	3.2 (1.3)	<0.0001
21. Made you feel conscious about the size and appearance of your stomach?	1.4 (0.8)	3.8 (1.3)	<0.0001
22. Made you concerned about soiling bed linen?	1.4 (0.8)	3.8 (1.3)	<0.0001
23. Made you feel sad, discouraged, or hopeless?	1.2 (0.5)	2.7 (1.2)	<0.0001
24. Made you feel down hearted and blue?	1.2 (0.6)	2.6 (1.2)	<0.0001
25. Made you feel wiped out?	1.4 (0.7)	3.4 (1.3)	<0.0001
26. Caused you to be concerned or worried about your health?	1.2 (0.6)	3.5 (1.1)	<0.0001
27. Caused you to plan activities more carefully?	1.2 (0.6)	3.4 (1.2)	<0.0001
28. Made you feel inconvenienced about always carrying extra pads, tampons, and clothing to avoid accidents?	1.3 (0.7)	3.8 (1.3)	<0.0001
29. Caused you embarrassment?	1.1 (0.3)	2.8 (1.3)	<0.0001
30. Made you feel uncertain about your future?	1.1 (0.5)	2.6 (1.3)	<0.0001
31. Made you feel irritable?	1.4 (0.8)	3.4 (1.2)	<0.0001
32. Made you concerned about soiling outer clothes?	1.3 (0.8)	3.7 (1.3)	<0.0001
33. Affected the size of clothing you wear during your periods?	1.2 (0.6)	3.3 (1.5)	<0.0001
34. Made you feel that you are not in control of your health?	1.1 (0.4)	3.1 (1.3)	<0.0001
35. Made you feel weak as if energy was drained from your body?	1.4 (0.8)	3.6 (1.2)	<0.0001
36. Diminished your sexual desire?	1.3 (0.8)	3.4 (1.3)	<0.0001
37. Caused you to avoid sexual relations?	1.6 (1.2)	3.2 (1.4)	<0.0001

FTG, fibroid treatment group; NCG, normal control group; SD, standard deviation; UFS-QOL, uterine fibroid symptom and quality-of-life questionnaire.

Table 3 – UFS-QOL and SF-36 scores at baseline: NCG versus FTG.

	Mean (SD)		*P* value
	NCG (N = 98)	FTG (N = 274)	
UFS-QOL			
Symptom severity*	15.3 (14.5)	64.8 (20.0)	<0.0001
Concern†	91.4 (16.8)	31.5 (27.9)	<0.0001
Activities†	95.2 (10.8)	42.9 (26.2)	<0.0001
Energy/mood†	90.9 (15.4)	43.5 (25.0)	<0.0001
Control†	95.9 (8.9)	45.8 (26.4)	<0.0001
Self-consciousness†	92.1 (15.1)	36.2 (29.3)	<0.0001
Sexual function†	88.2 (24.6)	42.9 (31.9)	<0.0001
HRQL total†	92.8 (11.7)	40.8 (22.1)	<0.0001
SF-36†			
Physical functioning	93.2 (15.5)	75.1 (26.0)	<0.0001
Role physical	93.4 (19.8)	47.9 (41.9)	<0.0001
Bodily pain	86.0 (16.1)	48.5 (24.3)	<0.0001
General health	83.3 (13.4)	67.5 (20.4)	<0.0001
Vitality	62.8 (18.0)	36.1 (21.5)	<0.0001
Social functioning	90.3 (16.4)	62.4 (28.1)	<0.0001
Role emotional	89.3 (23.8)	57.2 (42.8)	<0.0001
Mental health	78.8 (12.9)	62.9 (19.8)	<0.0001
Physical component score	54.7 (5.9)	43.1 (9.6)	<0.0001
Mental component score	51.3 (7.2)	41.6 (11.3)	<0.0001

FTG, fibroid treatment group; HRQL, health-related quality of life; NCG, normal control group; SD, standard deviation; SF-36, short form 36 health survey; UFS-QOL, uterine fibroid symptom and quality-of-life questionnaire.

* Scores range from 0 to 100; higher scores indicate greater symptom severity.

† Scores range from 0 to 100; higher scores indicate better HRQL.

sistency (Table 4). Internal consistency reliability was also very good at the 6-month follow-up (range: 0.77–0.99) and 1-year follow-up visits (range: 0.78–0.97).

Spearman's correlations were used to assess the concurrent validity of the UFS-QOL with the SF-36. Correlations among all UFS-QOL HRQL subscales and the SF-36 ranged from −0.04 (concern-role physical) to 0.40 (energy/mood-general health and control-physical component score) in the NCG and from 0.15 (self-consciousness-general health) to 0.66 (control-social functioning) in the FTG. Although the subscales on the UFS-QOL and SF-36 measure different domains, specific subscales of each scale are better suited to evaluate concurrent validity. For example, the UFS-QOL symptom severity subscale was moderately correlated with the SF-36 bodily pain subscale (NCG: r = –0.39; FTG: r = –0.43). Among women in the FTG, the symptom severity subscale was also moderately correlated with each of the SF-36 subscales.

Each UFS-QOL subscale was responsive to changes in the FTG condition with no significant change observed among the NCG over time. As expected, there were no significant changes in any of UFS-QOL subscale scores in the NCG from baseline to 6-month follow-up. Mean change scores ranged from 0.7 (concern) to 2.9 (energy/mood), with effect sizes ranging from 0.04 to 0.19 (Table 5). The UFS-QOL was highly responsive to treatment-related changes among the FTG. In the FTG, the mean change score from baseline at the 6-month follow-up for symptom severity was −43.7 ($P < 0.001$) representing an effect size of −2.18 and an SEM of 7.45 (Table 5). Mean change scores for the HRQL subscales ranged from 35.3 (sexual function) to 50.8 (concern) with effect sizes between 1.1 and 1.8. At the 1-year follow-up, the pattern of responsiveness results was similar to that observed at the 6-month follow-up. Again, the NCG indicated no significant changes in any of the scales, whereas the women in the FTG reported significant improvements in each of the scales (all $P < 0.001$). When examining the cumulative distribution curves at 6 months for the FTG, 90% of women who reported that they were doing "much better" on an overall treatment effect scale also had a 20 point or greater improvement in the symptom severity scale versus 70% of women reporting that they were "a little better" and 60% of women who reported they were "no better or worse" (Fig. 1). Similar results were noted at 1 year.

Twenty-nine women in the NCG and 59 women in the FTG completed their retest questionnaires within 2 weeks of the 6-month follow-up visit. Of the 59 FTG women, only 8 women reported being stable or unchanged during the 2-week retest period, therefore the test–retest analysis could not be conducted within this group. Test–retest reliability, however, was assessed in the NCG group among the 26 normal women who reported to be stable. The scores on the UFS-QOL subscales were largely stable with change scores ranging between 0.6 (self-consciousness) and 8.5 (sexual function) (Table 6). A statistically significant change was present in the energy/mood subscale and HRQL total; however, the change score for these two scales was 2.3 on a 0 to 100 scale, indicating very little change had occurred. The Spearman correlations for the subscales ranged from 0.37 for the control subscale to 0.84 for the sexual function subscale (all $P < 0.05$, except Control which was not significant) (Table 6). The low correlations noted for both symptom severity (0.42) and control (0.37) are likely due to lack of variability in symptoms in this sample of normal control women.

The CFAs were performed using the uterine fibroid group data; normal controls were excluded from the CFA. The CFA confirmed the factor structure of the UFS-QOL subscales and indicate that the UFS-QOL subscales of concern, activities, energy/mood, self-conscious, control, and sexual function fit the data very well with Bentler's CFI and all exceeded 0.90 (range: activities CFI = 0.92 to sexual functioning and self-conscious CFI = 0.99). The symptom severity and energy subscales also fit the data well with CFIs of 0.99 and 0.96, respectively, however, only after hierarchical latent models were used to improve fit.

Discussion

Assessing the impact of uterine fibroids in women has been an ongoing challenge for researchers and clinicians. Women with

Table 4 – Internal consistency reliability (Cronbach's alpha) of the UFS-QOL at baseline.

UFS-QOL	Number of items (*k*)	Cronbach's alpha	
		NCG (N = 98)	FTG (N = 274)
Symptom severity	8	0.82	0.79
Concern	5	0.90	0.92
Activities	7	0.90	0.92
Energy/mood	7	0.92	0.92
Control	5	0.73	0.89
Self-consciousness	3	0.87	0.80
Sexual function	2	0.84	0.91
HRQL total	29	0.95	0.96

FTG, fibroid treatment group; HRQL, health-related quality of life; NCG, normal control group; UFS-QOL, uterine fibroid symptom and quality-of-life questionnaire.

Table 5 – Responsiveness of the UFS-QOL.

UFS-QOL scale	Normal controls			Fibroid treatment group		
	Mean change score*	P value†	Effect size‡	Mean change score*	P value†	Effect size‡
6 Months						
Symptom severity§	−1.5	0.3654	−0.11	−43.7	<0.0001	−2.18
Concern‖	0.7	0.5811	0.04	50.8	<0.0001	1.81
Activities‖	1.2	0.2708	0.12	43.2	<0.0001	1.66
Energy/mood‖	2.9	0.0510	0.19	39.0	<0.0001	1.58
Control‖	1.5	0.1136	0.16	42.0	<0.0001	1.61
Self-consciousness‖	2.1	0.1371	0.15	42.4	<0.0001	1.44
Sexual function‖	1.8	0.3427	0.08	35.3	<0.0001	1.11
HRQL total‖	1.7	0.1113	0.15	42.7	<0.0001	1.96
12 Months						
Symptom severity§	−1.0	0.5218	−0.07	−47.2	<0.0001	−2.35
Concern‖	2.0	0.1962	0.12	54.8	<0.0001	1.92
Activities‖	0.9	0.4290	0.08	46.1	<0.0001	1.77
Energy/mood‖	0.7	0.6328	0.05	41.5	<0.0001	1.65
Control‖	1.4	0.2066	0.15	44.4	<0.0001	1.69
Self-consciousness‖	2.7	0.0633	0.18	42.4	<0.0001	1.43
Sexual function‖	2.3	0.3654	0.09	36.4	<0.0001	1.14
HRQL total‖	1.4	0.2353	0.12	45.2	<0.0001	2.03

HRQL, health-related quality of life; UFS-QOL, uterine fibroid symptom and quality-of-life questionnaire.

* Calculated as 6-month visit score minus baseline score; scores range from −100 to 100.

† Paired *t* tests comparing responses at baseline and 6 months.

‡ Calculated as score difference divided by standard deviation of baseline score.

§ Scores range from 0 to 100; higher scores indicate greater symptom severity.

‖ Scores range from 0 to 100; higher scores indicate better HRQL.

uterine fibroids experience a wide range of symptoms, including heavy bleeding, menstrual cramps, and urinary frequency and urgency that are also present to some degree among normal, menstruating women. Although menstruating women have menstrual symptoms that are captured by the UFS-QOL, the UFS-QOL clearly differentiates between women with and without fibroids, highlighting the profound impact that uterine fibroids and associated symptoms have on HRQL. Additionally, this study documented the responsiveness of the UFS-QOL to fibroid treatment and also documented the stability of the UFS-QOL in normal patients.

The primary limitation in clinical research of uterine fibroids has been the lack of outcome measures of the symptoms and associated condition-specific HRQL experienced by two important groups, normal women in the same age range as fibroid patients and women who undergo hysterectomy. Without data from these benchmark groups, it is difficult to assess the comparative effectiveness of other interventional or pharmacologic fibroid therapies. The ability to compare the outcomes of uterus-sparing fibroid therapies with hysterectomy advances the use of patient-reported outcomes in the field of fibroid therapy. Given that other studies evaluating newer fibroid therapies have also used the UFS-QOL

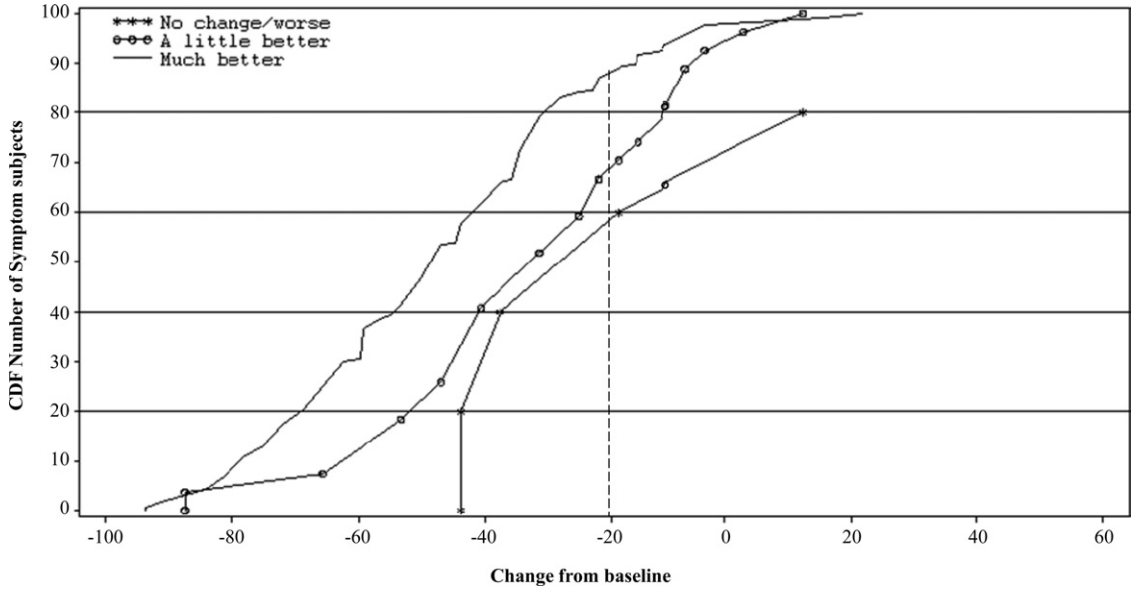

Fig. 1 – The uterine fibroid symptom and quality of life questionnaire symptom severity score change at month 6 in the fibroid treatment group.

Table 6 – Test–retest reliability of the UFS-QOL: normal control group.

UFS-QOL scale	6-month visit score N = 26 (mean, SD)	Retest visit score N = 26 (mean, SD)	Mean change score*	P value†	Intraclass correlation	Spearman correlation‡
Symptom severity§	14.8 (14.0)	12.4 (12.8)	−2.4	0.2734	0.66	0.42^
Concern‖	90.4 (22.4)	92.7 (17.6)	2.3	0.1034	0.94	0.74^^^
Activities‖	95.6 (11.0)	96.3 (9.5)	0.7	0.3269	0.94	0.78^^^
Energy/mood‖	92.7 (13.2)	95.1 (12.3)	2.3	0.0380	0.90	0.81^^^
Control‖	96.7 (7.9)	99.2 (2.3)	2.5	0.0618	0.34	0.37
Self-consciousness‖	96.2 (9.5)	96.8 (9.5)	0.6	0.7029	0.61	0.59^^
Sexual function‖	85.1 (28.5)	93.0 (12.5)	8.5	0.0604	0.49	0.84^^^
HRQL total‖	93.5 (11.8)	95.6 (9.8)	2.3	0.0145	0.90	0.83^^^

HRQL, health-related quality of life; UFS-QOL, uterine fibroid symptom and quality-of-life questionnaire.

* Calculated as 6-month visit score minus baseline score; scores range from −100 to 100.

† Paired *t* tests comparing responses at 6 months and retest visits.

‡ ^$P < 0.05$; ^^$P < 0.01$; ^^^$P < 0.001$.

§ Scores range from 0 to 100; higher scores indicate greater symptom severity.

‖ Scores range from 0 to 100; higher scores indicate better HRQL.

[7,8,10], having benchmark values among normal women as well as post-hysterectomy women provides context for cross-study evaluation as well as a common outcome tool that can be used for future comparative effectiveness research. It should be noted that the UFS-QOL offers more information than the traditional pictorial blood loss assessment diary [17], which has been used in many fibroid studies as a measure of menorrhagia. The pictorial blood loss assessment only captures information regarding blood loss without assessing other fibroid symptoms (e.g., pain or bulk symptoms) or the impact of fibroid symptoms on HRQL.

The initial psychometric validation study of the UFS-QOL [5] included an item reduction component to reduce the UFS-QOL from 72 to 37 items. This is the first study to assess the validity, reliability, and responsiveness of the final 37-item version of the UFS-QOL. The CFA performed in this study confirmed the factor structure of the final 37-item version of the UFS-QOL subscales. In addition, only 29 normal controls and 110 women with uterine fibroids were enrolled in the initial psychometric validation study. A larger cohort of women with and without uterine fibroids was enrolled in this study, which allowed for a more robust analysis of validity and reliability, and an analysis of the responsiveness of the UFS-QOL to treatment.

Although the UFS-QOL is highly responsive to the treatments in this study, it is difficult to recommend a minimal important difference (MID) that would define "responders" per se from this interventional group. First, this analysis is limited in that only 5 women in the FTG reported that their condition was worse after treatment whereas the vast majority of women experienced large improvements in their UFS-QOL scores. Secondly, 90% of women who responded that they were "much better," reported a 20 point or greater reduction in their symptom severity score of the UFS-QOL. This overwhelming large positive response among women with interventional treatments for their fibroids makes estimating an MID problematic. An MID recommendation from this study would likely be too high for pharmacologic treatments or other non-interventional treatments.

The authors acknowledge the limitations of this study. This study utilized a convenience sample of women with and without uterine fibroids who were enrolled as they were scheduled for uterine fibroid treatment (FTG) or presented to the office (NCG). Treatment assignment was not randomized or blinded; in fact, the women in this study not only knew but actively selected the treatment they received, which may have produced a bias in treatment outcome. Because scores across the three fibroid treatment groups were similar [11], however, this is likely not the case. In addition, the test–retest analyses were limited by the small sample of women who reported being stable or unchanged during the 2-week retest period. The sample size in the FTG was not adequate to perform these analyses so the analyses were only conducted within the NCG. Further evaluation of the test–retest reliability of the UFS-QOL in women with fibroids who are not undergoing treatment is warranted. Lastly, this study had been designed for 5 years of follow-up to provide long-term data regarding each treatment group. Unfortunately, with the loss of funding, the long-term data were not captured and are still needed to evaluate long-term outcomes of these fibroid treatments.

Despite these limitations, this study demonstrates the validity and reliability of the UFS-QOL and the responsiveness of the UFS-QOL to therapeutic change. The UFS-QOL is a valid and reliable measure that can be utilized to assess and compare outcomes of uterine fibroid therapies.

Source of financial support: This study was supported by a grant from the US Army (USARMC Contract Number A-12818.2b Gynecologic Disease Program).

REFERENCES

[1] Cramer SF, Patel A. The frequency of uterine leiomyomas. Am J Clin Pathol 1990;94:435–8.
[2] Pron G, Cohen M, Soucie J, et al. The Ontario uterine fibroid embolization trial. Part 1. Baseline patient characteristics, fibroid burden, and impact on life. Fertil Steril 2003;79:112–9.
[3] Wegienka G, Baird DD, Hertz-Picciotto I, et al. Self-reported heavy bleeding associated with uterine leiomyomata. Obstet Gynecol 2003; 101:431–7.
[4] Lippman SA, Warner M, Samuels S, et al. Uterine fibroids and gynecologic pain symptoms in a population-based study. Fertil Steril 2003;80:1488–94.
[5] 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.
[6] 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.
[7] Fiscella K, Eisinger SH, Meldrum S, et al. Effect of mifepristone for symptomatic leiomyomata on quality of life and uterine size: a randomized controlled trial. Obstet Gynecol 2006;108:1381–7.
[8] Goodwin SC, Spies JB, Worthington-Kirsch R, et al. Uterine artery embolization for treatment of leiomyomata: long-term outcomes from the FIBROID Registry. Obstet Gynecol 2008;111:22–33.

[9] Smith WJ, Upton E, Shuster EJ, et al. Patient satisfaction and disease specific quality of life after uterine artery embolization. Am J Obstet Gynecol 2004;190:1697–703;discussion 703–6.
[10] Stewart EA, Gostout B, Rabinovici J, et al. Sustained relief of leiomyoma symptoms by using focused ultrasound surgery. Obstet Gynecol 2007;110:279–87.
[11] Spies JB, Bradley LD, Guido R, et al. Outcomes from leiomyoma therapies: comparison with normal controls. Obstet Gynecol 2010; 116:641–52.
[12] Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30:473–83.
[13] Nunnally JC, Bernstein IH. Psychometric Theory. 3rd ed. New York: McGraw-Hill, 1994.
[14] Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002;11:193–205.
[15] Cohen JD. Statistical Power Analysis for the Behavioral Sciences (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
[16] Bollen KA. Structural Equations with Latent Variables. New York, New York: John Wiley & Sons Inc., 1989.
[17] 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.