American Journal of Obstetrics and Gynecology (2005) **193**, 103–13

ELSEVIER

American Journal of OBSTETRICS and GYNECOLOGY

www.ajog.org

Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7)

M. D. Barber, MD, MHS,[a,*] M. D. Walters, MD,[a] R. C. Bump, MD[b]

Department of Obstetrics and Gynecology at the Cleveland Clinic Foundation, Cleveland, OH[a] and Lilly Research Labs, Indianapolis, IN[b]

Received for publication August 13, 2004; revised November 29, 2004; accepted December 14, 2004

KEY WORDS
Quality of life
Urinary incontinence
Fecal incontinence
Pelvic organ prolapse
Pelvic floor

Objective: To develop short forms of 2 valid and reliable condition-specific quality-of-life questionnaires for women with disorders of the pelvic floor including urinary incontinence, pelvic organ prolapse, and fecal incontinence (Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire).

Study design: Data from the 100 women who contributed to the development and validation of the Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire long forms were used to develop the short-form questionnaires. All subsets regression analysis was used to find the items in each scale that best predicted the scale score on the respective long form. When different items appeared equivalent, a choice was made on item content. After development, the short forms and the Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire long forms were administered preoperatively to 45 women with pelvic floor disorders scheduled to undergo surgery to evaluate the correlation between short and long forms in a second independent population. The short forms were readministered 3 to 6 months postoperatively to assess the responsiveness of the instruments.

Results: The short-form version of the Pelvic Floor Distress Inventory has a total of 20 questions and 3 scales (Urinary Distress Inventory, Pelvic Organ Prolapse Distress Inventory, and Colorectal-Anal Distress Inventory). Each short-form scale demonstrates significant correlation with their long-form scales ($r = .86$, $r = .92$, and $r = .93$, respectively, $P < .0001$). For the Pelvic Floor Impact Questionnaire short form, the previously developed short form for the Incontinence Impact Questionnaire-7 was used as a template. The 7 items identified in the previously developed Incontinence Impact Questionnaire-7 short form correlate highly with the Incontinence Impact Questionnaire long form ($r = .96$, $P < .0001$) as well as the long forms of the Colorectal-Anal Impact Questionnaire scale ($r = .96$, $P < .0001$) and the Pelvic Organ Prolapse Impact Questionnaire ($r = .94$, $P < .0001$). All subsets regression analysis did not identify any items or combination of items that correlated substantially better for any of the 3 scales. The scales of the Pelvic Floor Distress Inventory-20 and Pelvic Floor Impact Questionnaire-7 maintained their excellent correlation to the Pelvic Floor Distress Inventory and Pelvic Floor Impact

* Reprint requests: Matthew D. Barber, MD MHS, Cleveland Clinic Foundation, 9500 Euclid Avenue, Desk A81, Cleveland Ohio 44195.
E-mail: barberm2@ccf.org

doi:10.1016/j.ajog.2004.12.025

Questionnaire long forms in the second independent sample (r = .88 to .94 for scales of Pelvic Floor Distress Inventory-20; r = .95 to .96 for scales of Pelvic Floor Impact Questionnaire-7, $P < .0001$ for all). The test-retest reliability of each scale was good to excellent (intraclass correlation coefficient 0.70 to 0.93, $P < .001$ for all scales). The scales and summary scores of the Pelvic Floor Distress Inventory-20 and Pelvic Floor Impact Questionnaire-7 demonstrated moderate to excellent responsiveness 3 to 6 months after surgery.

Conclusion: The Pelvic Floor Distress Inventory-20 and Pelvic Floor Impact Questionnaire-7 are valid, reliable, and responsive short forms of 2 condition-specific quality-of-life questionnaires for women with pelvic floor disorders.

Disorders of the pelvic floor, such as urinary incontinence, fecal incontinence, and pelvic organ prolapse, can have a significant impact on the quality of a woman's life. It is, therefore, important to measure quality of life in women with pelvic floor disorders when evaluating the efficacy of a particular therapy or comparing symptom severity between patients or groups. In 2001, 2 condition-specific quality-of-life instruments were developed for women with all forms of pelvic floor disorders, the Pelvic Floor Distress Inventory (PFDI) and the Pelvic Floor Impact Questionnaire (PFIQ).[1] These two instruments are based on the structure and content of 2 widely used condition-specific quality-of-life questionnaires for women with lower urinary tract dysfunction, the Urinary Distress Inventory (UDI) and the Incontinence Impact Questionnaire (IIQ), which were originally described by Shumaker et al.[2] The PFDI and PFIQ together can be used by clinicians and researchers to measure the extent to which lower urinary tract, lower gastrointestinal tract, and pelvic organ prolapse symptoms affect the quality of life of women who suffer from disorders of the pelvic floor. Each has been shown to be psychometrically valid and reliable.[1] Despite the strengths of these 2 complementary questionnaires, their comprehensive nature and relative length may be inefficient or impractical for some clinical or research situations. To remedy this, we sought to develop short form versions of these 2 questionnaires that are valid, reliable, and responsive to change.

Material and methods

Item selection

Data from the 100 women who contributed to the development and validation of the PFDI and PFIQ long forms were used to develop the short form questionnaires (group 1). Specific details of this patient group and the long versions of the PFDI and PFIQ can be found elsewhere.[1] In brief, the PFDI serves the role of both a symptom inventory and a measure of the degree of bother and distress caused by the broad array of pelvic floor symptoms. It includes all of the items in the original UDI instrument plus items relating to pelvic organ prolapse and lower gastrointestinal dysfunction. The PFIQ was designed to assess life impact in women with pelvic floor disorders and like the PFDI contains all of the items included in the original IIQ as well as items related to other pelvic floor disorders. The PFDI consists of 46 questions separated into 3 scales, the UDI, Pelvic Organ Prolapse Distress Inventory (POPDI), and the Colorectal-Anal Distress Inventory (CRADI). Similarly, the PFIQ consists of 3 scales of 31 questions each, the Urinary Impact Questionnaire (UIQ), the Pelvic Organ Prolapse Impact Questionnaire (POPIQ), and the Colorectal-Anal Impact Questionnaire (CRAIQ). The PFDI and PFIQ have been shown to have good test-retest reliability (intraclass correlation coefficients [ICCs] .86 and .87, respectively) and excellent internal consistency (Cronbach's alpha .88 and .97). The scales of the PFDI and PFIQ demonstrate significant association with appropriate measures of symptom severity and pelvic floor diagnoses, thereby demonstrating construct validity.

To determine which items would be included in the short forms of the PFDI and PFIQ, all subsets regression analysis was used to find the items in each scale that best predicted the scale score on the long form. When different items appeared equivalent statistically, a choice was made on item content. The UDI-6 and IIQ-7 short forms have been previously validated and are widely used.[3] Therefore, for the lower urinary tract scales, we sought to determine whether any different items or item combinations were substantially better than these existing scales. If not, we planned to include these scales in our instruments. After the item content for the short forms was determined, a multidisciplinary expert panel that included urogynecologists, female urologists, a colorectal surgeon, a pelvic floor physical therapist, and a psychometrician reassessed face and content validity.

External validation

After selection of the items to be included in the PFDI and PFIQ short forms, external validation of instruments in a new patient population was performed. Forty-five women who presented to the urogynecology clinic at the Cleveland Clinic Foundation were prospectively enrolled (group 2). Each enrolled subject completed an institutional review board–approved informed

consent process. Subjects were included if they were age 18 years or older; had one or more pelvic floor disorders including urinary incontinence, voiding dysfunction, pelvic organ prolapse (stage 2 or greater), defecatory dysfunction, or rectal prolapse; and were scheduled for reconstructive pelvic surgery or continence surgery. Subjects were excluded if they were mentally incapable of completing the self-administered questionnaires. Preoperatively all subjects underwent a standardized evaluation that included a structured urogynecologic history, physical examination with pelvic organ prolapse quantitation[4] and administration of the Medical Outcomes Study Short Form 36 (SF-36)[5], a generic health-related quality-of-life questionnaire, as well as the short and long forms of the PFDI and PFIQ. The short and long versions were completed at least 1 week apart to minimize subject recall. Scale scores from the short form were compared with scale scores of the long form using Pearson's correlation coefficient. One-week test-retest reliability for the short-form scales and summary scores were assessed using the ICC, a standard method of assessing test-retest reliability of continuous data. Subjects completed the baseline questionnaires during their initial clinic visit. The 1-week retest questionnaires were completed at home and returned through the mail. Methods, definitions, and diagnostic criteria of pelvic floor disorders conform to the standards recommended by the International Continence Society and National Institutes of Health except where otherwise noted.[4,6,7]

Responsiveness

Responsiveness, or sensitivity to change, refers to an instrument's ability to detect change that occurs as the result of therapy or disease progression. Adequate responsiveness is an essential property for any questionnaire intended to evaluate the effect of a treatment. To assess the responsiveness of the PFDI and PFIQ short forms, subjects in group 2 completed the short forms and the SF-36 again 3 to 6 months after surgery. These were completed at home and returned through the mail. A variety of statistics have been used to assess responsiveness and no single one has proven to be superior. We, therefore, chose to evaluate responsiveness of the PFDI and PFIQ short forms using several different techniques. First, preoperative scores were compared with postoperative scores using the paired *t* test. Next, two commonly used measures of responsiveness were assessed: effect size,[8] which is the change in mean score divided by the SD of the baseline, and standardized response mean (SRM),[9] which is the change in mean scores over the SD of the change. For both statistics a value of 0.5 to 0.7 is considered moderate responsiveness, 0.80 to 1.0 is considered good, and more than 1.0 is considered excellent.[10] These statistics were also computed for the scales of the SF-36 to evaluate the relative responsiveness of this commonly used generic quality-of-life measure in women with pelvic floor disorders.

One important aspect of responsiveness is the ability of an instrument or scale to detect small but important changes. This has led to the concept of the minimum clinically important difference (MCID), the smallest change in score associated with a clinically meaningful change in quality of life.[10] We evaluated the MCID of the summary scores of the PFDI and PFIQ short forms by comparing the change in scores 3 to 6 months after treatment with a global index of improvement. For this global index, subjects were asked 3 to 6 months after surgery to complete the sentence "Compared with before surgery, I feel that I am _____" using a 7-point scale from "very much worse" to "very much better." We chose to use this global scale to determine clinically meaningful change rather than some other outcome because global ratings such as this have been shown to provide the single best measure of significance of change from the individual perspective.[10] They also take into account more information that may affect health-related quality of life than other methods for assessing clinically meaningful change. Furthermore, because of the heterogeneous diagnoses found in our patient population, no single standard measure of symptom severity, such as pad testing or bladder diary, would be applicable for all subjects. The change in the summary scores of the PFDI and PFIQ short forms after surgical treatment were compared with the responses of the global rating scale using the Kruskal-Wallis test. In addition, the Spearman correlation coefficient (a nonparametric correlation coefficient) was used to evaluate the degree of association between the change in summary scores of the short forms and the global rating scale of improvement. The MCID of the summary scores of the PFDI and PFIQ short forms was defined as the mean change in score of those subjects indicating that they were "a little better" on the global rating scale.

The final method used to assess responsiveness of the PFDI and PFIQ short forms was an evaluation of the scale's ability to discriminate between those patients who were improved after treatment with those who were not.[9] The 7 ordinal categories of the global scale of improvement were collapsed into a dichotomous outcome of "better" and "worse." Logistic regression was used to determine the area under the receiver operating characteristic curve (ROC curve), or c-statistic, of the ability of the change in summary score of the PFDI and PFIQ short forms to accurately predict whether a subject felt that they were better or worse after surgery. A c-statistic of 1 indicates a perfect test, whereas a c-statistic of 0.5 indicates a test with no discriminative ability.

Sample size

An a priori sample size calculation was performed for the external validation and responsiveness portions of

Table I Baseline characteristics of the two study populations

	Group 1 (n = 100)	Group 2 (n = 45)	*P* value
Age*	56 ± 15	59 ± 12	.23
Caucasian (%)	83%	95%	.30
BMI*	29 ± 7	27 ± 6	.10
Parity†	2 (0-5)	2 (1-7)	.50
Insurance status (%)			.001
Private/HMO	44	71	
Medicaid/Medicare	56	29	
Previous hysterectomy (%)	68	51	.08
Previous pelvic reconstructive surgery (%)	32	20	.19
Pelvic floor disorders (%)			
Urodynamic stress incontinence	30	40	.32
Detrusor overactivity	16	16	.95
Voiding dysfunction	24	40	.75
Pelvic organ prolapse (stage 3 or 4)	27	58	.0007
Fecal incontinence	13	11	.96
Defecatory dysfunction	28	40	.21
Rectal prolapse	4	6	.96
Number of pelvic floor disorders per patient (%)			< .0001
0‡	24	0	
1	44	34	
2	26	40	
3+	6	26	

* Mean ± SD.
† Median (range).
‡ In group 1, 24% of patients did not meet the study definition of any of the pelvic floor diagnoses listed above but all had some symptom of pelvic floor dysfunction. In contrast, all patients in group 2 met the criteria for one or more pelvic floor disorders.

the study. Data from the validation study of the PFDI and PFIQ long forms were the basis for the sample size calculations. For the external validation portion of the study, a sample size of 45 subjects in group 2 provides greater than 80% power to detect a correlation coefficient of .90 or greater (the average correlation coefficient seen between the scales in the original study population), assuming a correlation coefficient of .75 or less between the scales of the short and long forms is undesirable (null hypothesis). In the assessment of responsiveness, a sample size of 45 has greater than 80% power to detect an effect size of 0.50 or greater using paired *t* test with .05 two-sided significance.

Results

The demographics and clinical diagnoses of the subjects in groups 1 and 2 are listed in Table I. The groups have similar age, racial distribution, parity, and body mass index. A greater proportion of subjects in group 1 had Medicaid or Medicare insurance than those in group 2. Also, subjects in group 2 were significantly more likely to have stage 3 or 4 prolapse and, in all, had greater number of pelvic floor disorders than those in group 1. The surgical procedures performed on subjects in group 2 are listed in Table II.

Table II Surgical procedures performed on group 2

Procedure	Percent
Hysterectomy	38
Trachelectomy	4
Anterior colporrhaphy	60
Posterior colporrhaphy	67
Paravaginal repair (abdominal or laparoscopic)	22
Vaginal vault suspension	55
Sacral colpopexy (abdominal or laparoscopic)	13
Sling procedure (including TVT)	40
Retropubic urethropexy (Burch)	20
Urethrolysis	2
Anal sphincteroplasty	4
Rectopexy	2

TVT, Tension-free vaginal tape.

Short-form questionnaires (PFDI-20 and PFIQ-7)

Table III compares the original PFDI and PFIQ with their short forms, the PFDI-20 and PFIQ-7. The short-form version of the PFDI has a total of 20 questions and 3 scales (UDI-6, POPDI-6, and CRADI-8) (see appendix). For the UDI scale, a short form already exists in the literature, the UDI-6.[3] Regression analysis did not identify a group or combination of questions substantially better than the UDI-6. ($r = .86$, $P < .0001$). Given this and the widespread use and familiarity of the UDI-6, a decision was made to incorporate the existing UDI-6 into the PFDI short form. For the POPDI scale, 6 items were identified for the short form. The combined score of these 6 questions demonstrated significant correlation with the POPDI long form (16 items) ($r = .92$, $P < .0001$). For the CRADI scale, 8 items that significantly correlated with the long form version of the scale (17 items) ($r = .93$, $P < .0001$) were identified. Each of the 3 scales of the PFDI-20 is scored from 0 (least distress) to 100 (greatest distress). The sum of the scores of these 3 scales serves as the overall summary score of the PFDI-20 and ranges from 0 to 300.

For the PFIQ short form, the previously developed short form for the IIQ was used as a template.[3] In the development of the PFIQ long version, each of the items in the long form of the IIQ were adapted to assess prolapse and lower gastrointestinal function as well as urinary function. The 7 items identified in the previously

Table III Comparison between the original PFDI and PFIQ and their short forms, the PFDI-20 and PFIQ-7

			Short form		
Original questionnaire	Original scales	No. of items	Questionnaire	Scales	No. of items
Pelvic Floor Distress Inventory (PFDI)		46	PFDI-20		20
	Urinary Distress Inventory (UDI)*	28		UDI-6‡	6
	Pelvic Organ Prolapse Distress Inventory (POPDI)	16		POPDI-6	6
	Colorectal-Anal Distress Inventory (CRADI)	17		CRADI-8	8
Pelvic Floor Impact Questionnaire (PFIQ)§		93	PFIQ-7‡		21
	Urinary Impact Questionnaire (UIQ)†	31		UIQ-7‡	7
	Pelvic Organ Prolapse Impact Questionnaire (POPIQ)	31		POPIQ-7‡	7
	Colorectal-Anal Impact Questionnaire (CRAIQ)	31		CRAIQ-7‡	7

* The UDI scale of the PFDI contains the 19 questions of the UDI described by Shumaker et al[2] and 9 additional questions.
† The UIQ scale of the PFIQ contains each of the 30 items in the IIQ described by Shumaker et al[2] and 1 additional question.
‡ The UDI-6 scale of the PFDI-20 short form contains the items of the previously validated instrument of the same name.[3]
§ Each scale of the PFIQ-7 is based on the content and structure of the previously validated IIQ-7.[3] Because the PFIQ-7 is intended for use in women with fecal incontinence as well as urinary incontinence, we have chosen to name the scale evaluating the impact of urinary dysfunction, the UIQ-7 rather than the original IIQ-7 to avoid confusion.

developed IIQ-7 short form correlate highly with the UIQ long form (r = .96, $P < .0001$) as well as the long forms of the CRAIQ scale (r = .96, $P < .0001$) and the POPIQ (r = .94, $P < .0001$). all subsets regression analysis did not identify any items or combination of items that correlated substantially better for any of the 3 scales. Again, given the familiarity and widespread use of the IIQ-7, a decision was made to use the appropriately adapted items from this scale and apply them to all 3 scales of the PFIQ short form. (see appendix) Additionally, using the same 7 items for each of the 3 scales in the PFIQ improves efficiency and discrimination among the impact of bowel, bladder, and urinary function. Like the PFDI-20, the PFIQ-7 has 3 scales that are scored from 0 (least impact) to 100 (greatest adverse impact) and an overall summary score (0 to 300)

The scales of the PFDI and PFIQ short forms maintained a high correlation with their long form counterparts when applied to a new population of 45 women with pelvic floor disorders (group 2; see Table IV). Additionally, each of the scales of the PFDI-20 and PFIQ-7 demonstrated good to excellent test-retest reliability with ICC values between .70 and .91.

Responsiveness

The responsiveness of the PFDI-20, PFIQ-7, and SF-36 can be seen in Table V. There was a significant improvement from baseline in each of the scales of the PFDI-20 short form and the overall summary score 3 to 6 months after surgery. Each scale demonstrated moderate to excellent responsiveness with effect size and SRM values ranging from .70 to 1.28. The sensitivity to change of the PFDI-20 as a whole was excellent with an effect size of 1.48 and SRM of 1.09. Similarly, the scales of the PFIQ short form also demonstrated a statistically significant improvement 3 to 6 months after surgery, but the responsiveness of these scales was somewhat less than that of the distress inventory short form. Overall, the PFIQ-7 short form demonstrated moderate responsiveness with an effect size of .67 and SRM of .63.

In contrast to the PFDI-20 and PFIQ-7, the majority of the scales of the SF-36 did not appear to be responsive to change in women with pelvic floor disorders undergoing surgery. Only the social functioning scale demonstrated a significant change from baseline, and it demonstrated only moderate responsiveness (effect size .75; SRM .57).

Overall, 7 subjects (16%) indicated that they were "a little better," 27 subjects (60%) indicated that they were "much better," and 8 subjects (17%) indicated that they were "very much better" on the global index of improvement 3 to 6 months after surgery. Only 3 subjects (7%) indicated that they were "a little worse," "much worse," or "very much worse," and no subjects indicated that they were the "unchanged." The summary scores of the PFDI-20 and PFIQ-7 demonstrated significant correlation with the subjects' global index of improvement (Table VI). A stepwise increase in change in score was seen for both instruments for each

Table IV Correlation of the scales of the short and long forms of the PFDI and PFIQ and test-retest reliability of the short forms*

	Correlation with long form scale (r)†	Test-retest reliability (ICC)‡
PFDI-20	—	.93
UDI-6	.86	.82
POPDI-6	.92	.91
CRADI-8	.93	.84
PFIQ-7	—	.77
UIQ-7	.96	.81
POPIQ-7	.94	.70
CRAIQ-7	.96	.81

* No correlation is shown for the summary scores of the PFDI-20 and PFIQ-7 because, unlike the short forms, the PFDI and PFIQ long forms do not have a summary score.

† $P < .0001$ for each.

‡ $P < .001$ for each.

level of improvement in the global index, with those indicating that they were "worse" demonstrating an increase in PFDI-20 and PFIQ-7 summary scores and those indicating they were "better" demonstrating a decrease in their summary scores. The mean decline in the summary score for subjects who indicated that they were "a little better" was 45 for the PFDI-20 and 36 for the PFIQ-7. The ability of the PFDI-20 and PFIQ-7 summary scores to successfully discriminate between subjects who indicated that they were "worse" after surgery from those who indicated that they were "better" was excellent with c-statistics of .95 and .88, respectively.

Comment

Pelvic floor disorders encompass a wide variety of interrelated clinical conditions that include urinary incontinence, fecal incontinence, pelvic organ prolapse, voiding dysfunction, and defecatory dysfunction that can adversely impact the lives of women. These disorders rarely result in mortality or severe morbidity; rather their primary impact is to adversely affect quality of life. Measuring quality of life is therefore essential when evaluating treatments and assessing the impact these disorders have on the lives of women. Although more than 14 instruments have been developed and validated for assessing the impact of urinary incontinence on the quality of life in women,[11] far fewer condition-specific quality-of-life instruments have been developed for fecal incontinence, and thus far, no instruments have been published that exclusively evaluate the impact of pelvic organ prolapse on quality of life. In 2001 we published the PFDI and the PFIQ, 2 reliable and valid condition-specific quality-of-life instruments intended for women with all forms of pelvic floor disorders.[1] The PFDI and PFIQ were designed to provide a comprehensive evaluation of the extent to which lower urinary tract, lower gastrointestinal tract, and pelvic organ prolapse symptoms affect the quality of life of women who suffer from disorders of the pelvic floor. In this study, short-form versions of the PFDI and PFIQ are presented, the PFDI-20 and PFIQ-7. These short forms demonstrate excellent correlation with the PFDI and PFIQ long forms and are reliable and responsive to change. Their relative length allows easy use in both the clinical and research setting.

Disorders of the pelvic floor share many common risk factors and frequently coexist.[12] Additionally, the treatment of one of these disorders can improve, worsen, or even predispose for another.[13-17] For example, the successful anatomic cure of a rectocele has been shown to improve bowel function in some patients but worsening it in others.[16,17] Similarly, Burch colposuspension has been demonstrated to be an effective treatment for stress urinary incontinence but can predispose patients to the development of pelvic organ prolapse, particularly enteroceles.[13,14] Because of these complex relationships, there are currently efforts to promote a comprehensive approach to the evaluation, treatment, and assessment of treatment outcomes in women with pelvic floor disorders.[7] A clear strength of the PFDI and PFIQ and their short forms is that they allow a comprehensive assessment of the effect of pelvic floor disorders on the quality of life of women, rather than assessing just one aspect of pelvic floor function such as urinary incontinence.

The PFDI and PFIQ long forms were based on the structure and content of 2 validated and widely used condition-specific quality-of-life questionnaires for women with urinary incontinence, the UDI and IIQ.[2] In fact, the PFDI and PFIQ contain within them all of the items of the UDI and IIQ as well as items evaluating the impact of lower gastrointestinal dysfunction and symptoms of pelvic organ prolapse on quality of life.[1] In keeping with this model, the PFDI-20 and PFIQ-7 include within them the UDI-6 and IIQ-7, the short-form versions of the UDI and IIQ.[3] The UDI-6 and IIQ-7 have been demonstrated to be valid, reliable, and responsive to change in several studies.[3,10] The World Health Organization's Second International Consultation on Incontinence rated the UDI and the UDI-6 among the 5 "highly recommended" questionnaires to assess symptoms of incontinence, and the IIQ and IIQ-7 among the 5 "highly recommended" questionnaires for assessing the impact of incontinence on quality of life.[11] Including the UDI-6 and IIQ-7 as scales within the PFDI-20 and PFIQ-7 allows clinicians and researchers to use these 2 valid, reliable, and popular questionnaires to assess the impact of the lower urinary tract on quality of life and simultaneously assessing the impact of any pelvic organ prolapse and/or bowel dysfunction.

Table V Mean change in scores and sensitivity to change of the PFDI-20, PFIQ-7, and SF-36

Scale	Pretreatment mean (SD) score	Post-treatment mean (SD) score	Mean change in score (SD)*	Effect size†	Standardized response mean (SRM)	Paired *t* test *P* value
PFDI-20	121.6 (48.2)	50.2 (38.9)	−71.4 (65.4)	1.48	1.09	< .0001
UDI-6	44.0 (23.2)	21.6 (21.7)	−22.4 (30.7)	.96	.73	< .0001
POPDI-6	46.6 (26.9)	12.3 (13.7)	−34.4 (29.8)	1.28	1.15	< .0001
CRADI-8	30.9 (19.0)	16.3 (14.7)	−14.6 (20.8)	.78	.70	< .0001
PFIQ-7	62.9 (58.3)	23.8 (39.4)	−39.1 (62.0)	.67	.63	< .001
IIQ-7	28.1 (22.1)	13.0 (18.1)	−15.0 (27.1)	.68	.55	.001
POPIQ-7	17.7 (23.9)	5.1 (16.7)	−12.6 (25.0)	.52	.50	.002
CRAIQ-7	17.2 (24.6)	5.7 (12.2)	−11.5 (22.7)	.47	.51	.002
SF-36						
Physical function	52.9 (26.4)	63.0 (25.1)	9.8 (32.0)	.37	.30	.16
Role, physical	57.3 (42.0)	53.2 (43.5)	−5.4 (45.2)	.12	.12	.57
Bodily pain	63.1 (26.2)	65.7 (25.7)	2.6 (34.2)	.10	.10	.71
General health	68.3 (18.5)	69.9 (15.9)	2.8 (26.9)	.15	.13	.53
Vitality	52.4 (15.8)	56.5 (15.8)	4.7 (15.8)	.30	.22	.29
Social functioning	71.9 (22.4)	88.0 (15.8)	16.8 (29.6)	.75	.57	.01
Role, emotional	77.8 (33.6)	85.5 (29.8)	7.3 (47.0)	.22	.15	.46
Mental health	76.8 (12.6)	78.8 (11.8)	2.6 (13.2)	.21	.20	.36
MCS	52.2 (8.1)	55.2 (9.1)	3.3 (11.9)	.41	.28	.19
PCS	39.4 (10.4)	40.7 (11.4)	1.2 (12.2)	.12	.10	.64

MCS, Mental component summary score; *PCS,* physical component summary scores.

* For the PFDI-20 and PFIQ-7, a negative change in score indicates improvement. For the SF-36, a positive change in score indicates improvement.

† Effect size is equal to the mean change in scores divided by the SD of the baseline score. SRM is equal to the mean change in scores divided by the SD of the change in scores.

Table VI Relationship between subject's global assessment of improvement and summary scores of the PFDI-20 and PFIQ-7

Global assessment of improvement category	Number (%)	Mean change in PFDI-20 (SD)	Mean change in PFIQ-7 (SD)
Worse*	3 (7%)	+22 (4)	+5 (25)
A little better	7 (16%)	−45 (74)	−36 (40)
Much better	27 (60%)	−73 (52)	−45 (57)
Very much better	8 (17%)	−106 (73)	−74 (46)
P value between categories		.02	.05
Spearman's rho (*P* value)		.41 (.005)	.37 (.02)
Area under ROC curve (c-statistic)†		.95	.88

* Combined category of subjects who indicated that they were "a little worse," "much worse," or "very much worse." No subjects indicated that they were "the same" after surgery.

† Area under the ROC curve (c-statistic) for the ability of the summary score to successfully discriminate between subjects who indicated that they were "worse" after surgery from those who indicated that they were "better."

Because the PFIQ-7 is intended for use in women with fecal incontinence as well as urinary incontinence, we have chosen to name the scale evaluating the impact of urinary dysfunction on quality of life, the UIQ-7 rather than the original IIQ-7 to avoid confusion. The components of the UIQ-7 scale are identical to the IIQ-7 described by Uebersax et al[3] in 1995, however.

Responsiveness, or sensitivity to change, is the ability of an instrument to detect a small but clinically important change.[9,10] It is an important psychometric property for any measure intended to assess an intervention. The responsiveness of a measure can have substantial impact on the interpretation of a clinical trial; however, this psychometric property is often neglected in the literature.[10,11,18] Use of a measure that is has poor responsiveness, or is "unresponsive," will increase the risk a type II error (assuming no difference when a difference in fact exists) and underestimate of the effect of treatment.[19] Many different methods have been described for measuring the responsiveness of an instrument without a clear consensus for which method is best.[9,10] We therefore chose to assess the responsiveness of the PFDI-20 and PFIQ-7 using several different methods. Three commonly used methods were used to

assess the responsiveness of the each of the 3 scales of the PFDI-20 short form and PFIQ-7 short form as well as their summary scores: paired *t* test, effect size, and SRM. Each method confirmed moderate to excellent responsiveness of the PFDI-20 and PFIQ-7 scales and summary scores. The UDI-6 and IIQ-7 have both demonstrated adequate responsiveness previously.[3,10] In contrast, Fitzgerald et al[20] found that the UDI-6 was responsive to objective change in continence status after continence surgery, whereas the IIQ-7 was not. Although we found adequate responsiveness for both the PFDI-20 and PFIQ-7, our findings do suggest that the scales of the PFDI-20 are more responsive than the scales of the PFIQ-7. Both instruments demonstrate an excellent ability to discriminate between improved and unimproved patients, however. The relative responsiveness of the long version of the PFDI and PFIQ, compared with their corresponding short forms, was not evaluated in this study and is currently unknown.

In contrast to the PFDI-20 and PFIQ-7, the SF-36 is relatively unresponsive to change in women undergoing surgery for pelvic floor disorders. The SF-36 is perhaps the most widely used generic quality-of-life questionnaire in the United States. It has been used in studies of urinary incontinence and fecal incontinence.[11,21] In general, generic quality-of-life instruments are less responsive than condition-specific quality-of-life instruments, and several studies have demonstrated poor responsiveness of the SF-36 in patients undergoing treatment for urinary incontinence.[11,18] Generic quality-of-life instruments do, however, have the distinct advantage of allowing comparisons across different groups and disease processes. Therefore, in studies evaluating therapies for pelvic floor disorders, using both a generic quality-of-life instrument such as the SF-36 and condition-specific instruments with demonstrated responsiveness like the PFDI-20 and PFIQ-7 seems prudent to ensure breadth, comparability across populations, and an adequate assessment of treatment effect.[18]

The smallest change in score associated with a clinically meaningful change in quality of life has been called the MCID.[10] The concept of MCID is important because small numerical differences in mean quality-of-life scores might give statistically significant results when large sample sizes are used, but statistical significance is not equivalent to clinical significance.[22] We attempted to determine the MCID of the PFDI-20 and PFIQ-7 summary scores by comparing these scores with a patient-assessed global index of improvement. Use of a global index such as this is one of the most commonly used approaches for establishing clinically meaningful change.[10] Global ratings of change have been shown to provide the single best measure of significance of change from the individual perspective.[10] A priori, we defined the MCID of the PFDI-20 and PFIQ-7 as the mean change in score of those patients who indicated that they were "a little better" after surgery. Based on this definition, a change of 45 points (15%) or more in the summary score of the PFDI-20 and a change of 36 points (12%) or more in the summary score of the PFIQ-7 would be considered as "clinically important" in patients undergoing surgery for pelvic floor dysfunction. Changes in score less than this may not be clinically important, even if statistically significant. However, given the relatively small sample size in this study and the wide variability of change in scores, these values need to be confirmed by larger studies. It should be noted that these values represent the within-treatment MCID. The minimum difference in scores between 2 treatment groups in a clinical trial that would be considered clinically significant, or the between-treatment MCID, was not evaluated in this study. Additionally, we did not evaluate the MCID of the individual scales of the PFDI-20 and PFIQ-7 in this study, only the summary scores. Future studies are warranted to determine the between-treatment MCID and the MCID of the individual PFDI-20 and PFIQ-7 scales.

Choosing between the short and long versions of the PFDI and PFIQ will depend on the intended use. The PFDI-20 and PFIQ-7 will be useful in clinical practice, research studies that include multiple questionnaires, and any situation in which there is a desire to minimize respondent burden and cost. When there is a need for detail on specific symptoms or individual domains of distress and life impact, the long-form versions may be more appropriate.

References

1. Barber MD, Kuchibhatla MN, Pieper CF, Bump RC. Psychometric evaluation of 2 comprehensive condition-specific quality of life instruments for women with pelvic floor disorders. Am J Obstet Gynecol 2001;185:1388-95.
2. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Qual Life Res 1994;3:291-306.
3. Uebersax JS, Wyman JF, Shumaker SA, McClish DK, Fantl JA. Short forms to assess life quality and symptom distress for urinary incontinence in women. Neurourol Urodyn 1995;14:131-9.
4. Bump RC, Mattiasson A, Bo K, Brubaker LP, De Laney JO, Klarskov P, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. Am J Obstet Gynecol 1996;175:10-7.
5. Ware JE. SF-36 heath survey: manual and interpretation guide. Boston: Health Institute, New England Medical Center; 1993.
6. Abrams P, Blaivas JG, Stanton SL, Andersen JT. The standardisation of terminology of lower urinary tract function. The International Continence Society Committee on Standardisation of Terminology. Scand J Urol Nephrol Suppl 1988;114:5-19.
7. Weber AM, Abrans P, Brubaker L, Cundiff G, Davis G, Dmochowski RR, et al. The standardization of terminology for researchers in female pelvic floor disorders. Int Urogynecol J Pelvic Floor Dysfunct 2001;12:178-86.

8. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:178-89.
9. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. Control Clin Trials 1991;12:142s-58s.
10. Crosby RD, Kolotkin RL, Rhys Willisams G. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395-407.
11. Donavan JL, Badia X, Corcos J, et al. Symptom and quality of life assessment. In: Abrams P, Cardozo L, Khoury S, Wein A, editors. Second International Consultation on Incontinence. Plymouth, UK: Health Publications; 2002.
12. Bump RC, Norton PA. Epidemiology and natural history of pelvic floor dysfunction. Obstet Gynecol Clin North Am 1998;25:723-46.
13. Kjolhede P, Noren B, Ryden G. Prediction of genital prolapse after Burch colposuspension. Acta Obstet Gynecol Scand 1996;75:849-54.
14. Wiskind AK, Creighton SM, Stanton SL. The incidence of genital prolapse after Burch colposuspension. Am J Obstet Gynecol 1992;197:399-404.
15. Bump RC, Hurt WG, Theofastous JP, Addison WA, Fantl JA, Wyman JF, et al. Randomized prospective comparison of needle colposuspension versus endopelvic fascia plication for potential stress incontinence prophylaxis in women undergoing vaginal reconstruction for stage III or IV pelvic organ prolapse. Am J Obstet Gynecol 1996;175:326-33.
16. Kahn MA, Stanton SL. Posterior colporrhaphy: its effects on bowel and sexual function. BJOG 1997;104:82-6.
17. Cundiff GW, Weidner AC, Visco AG, Addison WA, Bump RC. An anatomic and functional assessment of discrete defect rectocele repair. Am J Obstet Gynecol 1998;179:1451-6.
18. Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality of life instruments. J Clin Epidemiol 2003;56:52-60.
19. Testa MA, Nackley JF. Methods for quality of life studies. Ann Rev Public Health 1994;15:535-9.
20. FitzGerald MP, Kenton K, Shott S, Brubaker L. Responsiveness of quality of life measurements to change after reconstructive pelvic surgery. Am J Obstet Gyencol 2001;185:20-4.
21. Rockwood TH, Church JM, Fleshman JW, Kane RL, Mavrantonis C, Thorson AG, et al. Fecal incontinence quality of life scale: quality of life instrument for patients with fecal incontinence. Dis Colon Rectum 2000;43:9-16.
22. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality of life scores. J Clin Oncol 1998;16:139-44.

Appendix

Pelvic Floor Impact Questionnaire–short form 7

Instructions: Some women find that bladder, bowel, or vaginal symptoms affect their activities, relationships, and feelings. For each question place an **X** in the response that best describes how much your activities, relationships, or feelings have been affected by your bladder, bowel, or vaginal symptoms or conditions **over the last 3 months**. Please make sure you mare an answer in **all 3 columns** for each question.

How do symptoms or conditions relate to the following →→→→ usually affect your ↓	***Bladder or urine***	***Bowel or rectum***	***Vagina or pelvis***
1. Ability to do household chores (cooking, housecleaning, laundry)?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit
2. Ability to do physical activities such as walking, swimming, or other exercise?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit
3. Entertainment activities such as going to a movie or concert?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit
4. Ability to travel by car or bus for a distance greater than 30 minutes away from home?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit
5. Participating in social activities outside your home?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit
6. Emotional health (nervousness, depression, etc)?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit
7. Feeling frustrated?	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit	☐ Not at all ☐ Somewhat ☐ Moderately ☐ Quite a bit

Scoring the PFIQ – 7:
All of the items use the following response scale:
0, Not at all; 1, somewhat; 2, moderately; 3, quite a bit

Scales:
Urinary Impact Questionnaire (UIQ-7): 7 items under column heading “Bladder or urine.”
Colorectal-Anal Impact Questionnaire (CRAIQ-7): 7 items under column heading “Bowel or rectum.”
Pelvic Organ Prolapse Impact Questionnaire (POPIQ-7): 7 items under column heading “Pelvis or vagina.”

Scale scores: Obtain the mean value for all of the answered items within the corresponding scale (possible value 0 to 3) and then multiply by (100/3) to obtain the scale score (range 0 to 100). Missing items are dealt with by using the mean from answered items only.
PFIQ-7 Summary Score: Add the scores from the 3 scales together to obtain the summary score (range 0 to 300).

Pelvic Floor Distress Inventory–short form 20

Instructions: Please answer all of the questions in the following survey. These questions will ask you if you have certain bowel, bladder, or pelvic symptoms and, if you do, how much they bother you. Answer these by putting an **X** in the appropriate box or boxes. While answering these questions, please consider your symptoms over the **last 3 months.**

The PFDI-20 has 20 items and 3 scales.
All items use the following format with a response scale from 0 to 4.

Do you _____?

☐ No; ☐ Yes
0

If yes, how much does it bother you?
☐ 1 ☐ 2 ☐ 3 ☐ 4
Not at all Somewhat Moderately Quite a bit

Scales
Pelvic Organ Prolapse Distress Inventory 6 (POPDI-6):

1. Usually experience *pressure* in the lower abdomen?
2. Usually experience *heaviness or dullness* in the pelvic area?
3. Usually have a bulge or something falling out that you can see or feel in your vaginal area?
4. Ever have to push on the vagina or around the rectum to have or complete a bowel movement?
5. Usually experience a feeling of incomplete bladder emptying?
6. Ever have to push up on a bulge in the vaginal area with your fingers to start or complete urination?

Colorectal-Anal Distress Inventory 8 (CRADI-8):

7. Feel you need to strain too hard to have a bowel movement?
8. Feel you have not completely emptied your bowels at the end of a bowel movement?
9. Usually lose stool beyond your control if your stool is well formed?
10. Usually lose stool beyond your control if your stool is loose?
11. Usually lose gas from the rectum beyond your control?
12. Usually have pain when you pass your stool?
13. Experience a strong sense of urgency and have to rush to the bathroom to have a bowel movement?
14. Does part of your bowel ever pass through the rectum and bulge outside during or after a bowel movement?

Urinary Distress Inventory 6 (UDI-6):

15. Usually experience frequent urination?
16. Usually experience urine leakage associated with a feeling of urgency, that is, a strong sensation of needing to go to the bathroom?
17. Usually experience urine leakage related to coughing, sneezing, or laughing?
18. Usually experience small amounts of urine leakage (that is, drops)?
19. Usually experience difficulty emptying your bladder?
20. Usually experience *pain* or *discomfort* in the lower abdomen or genital region?

Scale scores: Obtain the mean value of all of the answered items within the corresponding scale (possible value 0 to 4) and then multiply by 25 to obtain the scale score (range 0 to 100). Missing items are dealt with by using the mean from answered items only.

PFDI –20 Summary Score: Add the scores from the 3 scales together to obtain the summary score (range 0 to 300).