

Validation of the Profile of Female Sexual Function (PF SF) in Surgically and Naturally Menopausal Women

LEONARD DEROGATIS

University of Maryland, Baltimore, Maryland, USA

JOHN RUST and SUSAN GOLOMBOK

City University, London, United Kingdom

CELINE BOUCHARD

Centre Médical Santé Femme, Québec, Canada

LILA NACHTIGALL

New York, New York, USA

CYNTHIA RODENBERG and JAMES KUZNICKI

Procter & Gamble Pharmaceuticals, Inc., Cincinnati, Ohio, USA

COLLEEN A. MCHORNEY

Indiana University School of Medicine and the Regenstrief Institute for Health Care, Indianapolis, Indiana, USA

The Profile of Female Sexual Function (PF SF) is a patient-based instrument for the measuring of loss of sexual function in menopausal women with low libido (hypoactive female sexual desire disorder). The instrument, which contains 37 items in seven domains (sexual desire, arousal, orgasm, sexual pleasure, sexual concerns, sexual responsiveness, and sexual self-image) and a single-item measure of overall satisfaction with sexuality, has been extensively developed and initially validated in over 500 oophorectomized women with low libido in North America, Europe, and Australia. Initial validation results showed the PF SF is capable of discriminating these patients from age-matched controls and produced consistent responses and sensitivity across geographies. The objective of this non-randomized, parallel-group study was to examine the psychometric

This work was funded by grants from Procter & Gamble Pharmaceuticals, Inc., Cincinnati, Ohio. The authors would like to thank Gloria Yiu for performing statistical analyses and Lisa Bosch for assistance in the preparation of the manuscript.

Address correspondence to Leonard Derogatis, University of Maryland, 655 Lombard Street, Baltimore, MD 21201-1579. E-mail: Derogati@son.umaryland.edu

properties of the final PFSF in an independent group of surgically menopausal women with low libido and to extend validation to naturally menopausal women with low libido. Participants from 16 study centers in North America included surgically (n = 59) and naturally (n = 88) menopausal women with low libido and their age-matched control subjects, both premenopausal (n = 57) and naturally menopausal (n = 47), who reported no problems with libido. Subjects completed the PFSF at baseline and again 4 weeks later. Adjusted mean scores for each of the seven domains were statistically significantly lower ($P < 0.0001$) in surgically menopausal women with low libido compared with age-matched control women, and in naturally menopausal women with low libido compared with naturally menopausal control women, demonstrating excellent discriminant validity. Test-retest reliability ranged from 0.57 to 0.91 for the seven domain scores, whereas internal-consistency reliability ranged from 0.74 to 0.95. Results of this research support the conclusion that the PFSF is a valid and reliable instrument for measurement of loss of sexual function in both naturally and surgically menopausal women with low libido.

Hypoactive sexual desire disorder (HSDD) is defined as a persistent deficiency or absence of sexual fantasies and desire for sexual activity that causes marked distress or interpersonal difficulty (American Psychiatric Association, 1994). Women who have undergone menopause, whether natural or surgical (following oophorectomy), may experience a significant reduction in sexual desire. Clinical evaluation of any intervention to alleviate this condition requires the availability of a psychometrically sound and appropriate outcomes measure for measurement of sexual desire and related aspects of sexual function. Review and evaluation of existing self-report measures indicated that they lacked the focused content and multinational linguistic validation specific to the population of menopausal women with HSDD. These deficiencies were perceived as particularly significant, given the high prevalence of HSDD relative to other sexual dysfunctions among women (Lauman, Paik, & Rosen, 1999).

The Profile of Female Sexual Function (PFSF) is a new, self-report, multinational instrument designed to measure sexual desire and associated symptoms in women with HSDD following menopause. Specifically, this instrument, which contains 37 items in seven separate domains and an overall sexual satisfaction question, has been developed to reflect the clinical phenomenology of this disorder and its effects on the patients' thoughts, feelings, behaviors, and emotions. Initial development of the PFSF was performed extensively in over 800 oophorectomized women with low libido across North America, Europe, and Australia. These studies have shown the PFSF to be

reliable and valid for use as an outcomes measure in oophorectomized women with hypoactive sexual desire disorder and to demonstrate robust psychometric properties across numerous geographies (McHorney et al., manuscript submitted). The objective of the current study was to extend and generalize the validation statement of the PFSF to an additional independent sample of surgically menopausal women with low libido and to a separate group of naturally menopausal women suffering from HSDD.

METHODS

Design and Subjects

This nonrandomized, parallel-group study included surgically and naturally menopausal women with low libido and age-matched controls. Prior to including naturally menopausal women with low libido in the present study, we conducted qualitative research in one-on-one interviews (15 women) and focus groups (two groups of 8 women each) in naturally menopausal women with low desire that indicated that the items of the PFSF had both face and content validity in this group. On the basis of menopausal status and libido, we enrolled women from 15 study centers in the United States and one in Canada in study groups. Inclusion and exclusion criteria consistent with the DSM-IV criteria for hypoactive sexual desire disorder (APA, 1994) were used to identify women with low libido. To comply with inclusion criteria, these women were required to report a satisfying sex life before menopause (either surgical or natural) with a meaningful loss in desire and decrease in sexual activity subsequent to menopause. In addition, subjects had to report personal distress or discomfort over their decreased level of desire for sexual activity. To qualify, women in the control groups had to characterize their sex life as good and satisfying, be satisfied with their current level of sexual activity, and express little or no concern with their level of interest in sex.

All women were 20 to 70 years old, in a stable, monogamous relationship for at least one year with a partner who was sexually functional and had a body mass index not less than 18 or greater than 30. Women in the surgically menopausal, low libido group were to have undergone bilateral salpingo-oophorectomy and total hysterectomy at least 1 year before study entry, to have been on a stable dose of estrogen therapy for at least 3 months, and to be willing to continue estrogen treatment while in the study. Women in the naturally menopausal, low libido group were to have had no spontaneous menstrual period for at least one year before entry, to have a uterus and at least one ovary, and to be on stable hormone therapy. Women in the control groups were to have at least one ovary and to be on stable hormone replacement therapy if postmenopausal.

Women were excluded from participation if they had chronic or acute life stress relating to major life change that might interfere with sexual activity.

Additionally, consistent with the DSM-IV criteria, we excluded women whose sexual dysfunction was thought to be due to a psychiatric disorder or was due to the use of a substance or other medical condition. Women who had dyspareunia, physical problems or previous sexual trauma that might interfere with sexual activity, who were depressed according to the Beck Depression Inventory-II (Beck, Steer, & Brown, 1996), or who were taking medications known to affect desire were excluded.

Women in the surgically menopausal, low libido group were age-matched to women in the control group. The control group consisted of premenopausal and naturally menopausal women. Women in the naturally menopausal, low libido group were age-matched to the subset of naturally menopausal control women.

Measurements

The PFSF was initially administered at baseline and 4 weeks later. To assess convergent validity, we also administered the Derogatis Interview for Sexual Function-Self Report (DISF-SR) (Derogatis, 1997) at baseline and 4 weeks. In addition, women answered a question regarding overall satisfaction with sexuality over the past 30 days, which they rated on a scale from one (poor) to five (excellent). All questionnaires were completed by the subjects in a private place after appropriate instruction, without assistance.

We computed a raw score for each of the domains of the PFSF (sexual desire, arousal, orgasm, sexual pleasure, sexual concerns, sexual responsiveness, and sexual self image) by summing the item scores, after appropriate reversal of positively worded items. Each raw score was then converted, via a linear transformation, to a 0 to 100 scale. Higher scores on all domains represent better sexual function.

Statistical Analyses

We assessed the psychometric properties of reliability and validity for each domain of the PFSF and the single-item measure of overall sexual satisfaction. Test-retest and internal consistency reliability were evaluated. We assessed test-retest reliability using an intraclass correlation coefficient (ICC; Shrout & Fleiss, 1979). The ICC measures the correlation between two responses from the same individual. Internal consistency of the domains was assessed by calculation of Cronbach's α coefficient. Test-retest and internal-consistency estimates were computed separately for each low libido group and for the control group.

We evaluated the instrument's discriminant and convergent validity to ascertain initial validation of the PFSF. The discriminant capacity of the instrument was assessed by two means. We assessed classical known-groups

validity (McHorney et al., 1993) by evaluating the ability of the instrument to discriminate between the mean responses of each low libido group and their respective control group using a repeated measures analysis. To account for variations in underlying baseline characteristics, we included age, referral source, marital status, and length of relationship with partner in the model. Additionally, we assessed the discriminant efficiency of the PFSF using the scale's ability to distinguish individual scores of women in the low libido group from those of the controls. More specifically, we used a receiver operating characteristics (ROC) analysis to establish the ability to correctly classify a subject's status (with or without low libido) based on their domain score. For each domain, the area under the ROC curve (AUC), an overall assessment of a tool's discriminative capability, was computed (Hanley & McNeil, 1982). Additionally, on the basis of preliminary cutpoints derived from previous studies, we estimated sensitivity, specificity, and positive and negative predictive values. These cutoff points were established in earlier validation studies of surgically menopausal women and age-matched controls (McHorney et al., manuscript submitted) and correspond to the values that classify correctly, as low libido or control, the maximum percentage of subjects in a population with a 50% prevalence rate of low libido. This amounts to treating incorrect classification of a woman without low libido as important as incorrect classification of a woman with low libido.

We evaluated convergent validity (that is, positive and substantial correlation with an instrument measuring a similar construct) by correlating domains of the PFSF and the DISF-SR. The DISF is an established, reliable, and valid gender-keyed brief outcomes measure, available in both self-report and semistructured interview formats, designed to provide a quantitative estimate of an individual's current sexual functioning (Derogatis, 1997). The DISF represents five primary domains of sexual functioning (sexual cognition/fantasy, sexual arousal, sexual behavior/experiences, orgasm, and sexual drive and relationship) and a summary or total score. The DISF differs from the PFSF principally in that the former is more of an omnibus scale, appropriate across a broad spectrum of sexual function/dysfunction, whereas the PFSF is specifically focused on the assessment of sexual desire and its consequences. In the convergent validation exercise, we prospectively identified prior to unblinding the study domains hypothesized to demonstrate a significant positive relationship. Although we expected moderate correlations between all domains of the two instruments, the PFSF domains and the DISF domains with which they were expected to be highly correlated were sexual desire with sexual cognition/fantasy, sexual arousal, drive and relationship; arousal with sexual arousal and orgasm; orgasm with sexual arousal and orgasm; and sexual responsiveness with sexual cognition/fantasy.

TABLE 1. Baseline Characteristics of the Subjects by Group

	Low libido surgical menopause <i>n</i> = 59	All control <i>n</i> = 104	Low libido natural menopause <i>n</i> = 88	Natural menopause control <i>n</i> = 47
Age (years)*	51.1 (1.02)	48.2 (0.74)	55.2 (0.57)	54.9 (0.73)
Race— <i>n</i> (%)				
Caucasian	54 (92)	96 (92)	83 (94)	45 (96)
Black	4 (7)	3 (3)	2 (2)	1 (2)
Hispanic	1 (2)	3 (3)	3 (3)	0
Other	0	2 (2)	0	1 (2)
Married— <i>n</i> (%)	48 (81)	67 (64)	74 (84)	29 (62)
Length of relationship (years)*	19.8 (1.66)	15.9 (1.19)	21.4 (1.42)	20.6 (1.99)

*Data presented as mean (SEM).

RESULTS

Subject Disposition and Baseline Characteristics

Of 251 subjects, 249 (99%) completed the study. Two subjects failed to complete the study; one was lost to follow-up and one voluntarily withdrew. Baseline characteristics of all subjects by enrollment group are shown in Table 1. Subjects in each of the low libido groups were similar to their respective controls in terms of age, height, weight, and body mass index (data not shown). Surgically menopausal subjects with low libido were more likely to be married to their partners and in relationships of longer duration than were the control subjects. Although naturally menopausal women with low libido were also more likely to be married than were the naturally menopausal control women, similar lengths of relationship were seen for the two groups.

Test-Retest Reliability

Test-retest reliability data are shown in Table 2. Overall, high to moderately high intraclass correlation coefficients were observed (ICC range from 0.61 to 0.91) for all domains except sexual concerns, indicating consistent responses and good temporal stability between the two test administrations, at baseline and at 4 weeks. Slightly lower but acceptable ICC values (ICC range from 0.57 to 0.74) were observed for the sexual concerns domain. Additionally, a slightly lower but acceptable ICC value was seen for the overall sexual satisfaction question (ICC range from 0.54 to 0.73). Such lower values are not unexpected for single-item measures.

Domain Representation and Item Consistency

A high degree of scale homogeneity was observed for all domains of the PFSF with Cronbach's alpha coefficients greater than or equal to 0.74 (Table 3).

TABLE 2. Reliability of PFSF: Test-Retest Correlations in Low Libido Surgically Menopausal Women, Low Libido Naturally Menopausal Women, and Controls

Domain	Test-Retest intraclass correlation		
	Low libido surgical menopause <i>n</i> = 59	Low libido natural menopause <i>n</i> = 88	All controls <i>n</i> = 104
Sexual desire	0.73	0.73	0.76
Arousal	0.61	0.62	0.68
Orgasm	0.71	0.76	0.82
Sexual pleasure	0.74	0.77	0.80
Sexual concerns	0.57	0.58	0.74
Sexual responsiveness	0.81	0.84	0.91
Sexual self image	0.62	0.67	0.78

By and large, this result tends to confirm the structural integrity of the PFSF domains and the achievement of well-integrated construct definitions.

Clinical Known-Groups Validity of the PFSF

The ability of the PFSF to discriminate between the comparison groups of interest is shown in Figure 1. For all domains, the mean scores for women with HSDD were significantly lower (i.e., reflecting poorer sexual function) than those of the age-matched control groups ($p < 0.001$). In addition, the mean scores across domains for the natural and surgical menopausal populations were very consistent, supporting an expectation derived from qualitative content validation that, regardless of the etiology of menopause, women with HSDD experience similar complaints and problems with sexual functioning. The mean scores (*SE*) for the satisfaction with sexuality question were significantly lower than their respective control groups for both

TABLE 3. Internal Consistency of PFSF: Cronbach's α in Low Libido Surgically Menopausal Women, Low Libido Naturally Menopausal Women, and Controls

Domain	Cronbach's α		
	Low libido surgical menopause <i>n</i> = 59	Low libido natural menopause <i>n</i> = 88	All controls <i>n</i> = 104
Sexual desire	0.94	0.89	0.93
Arousal	0.95	0.91	0.81
Orgasm	0.89	0.93	0.88
Sexual pleasure	0.96	0.94	0.93
Sexual concerns	0.91	0.83	0.74
Sexual responsiveness	0.93	0.91	0.89
Sexual self image	0.87	0.80	0.85

Note. Data from the first administration of the instrument.

the surgically menopausal women with low libido [1.65 (0.159) versus 3.77 (0.123); $P < 0.0001$] and naturally menopausal women with low libido [1.68 (0.130) versus 3.82 (0.127); $P < 0.0001$].

Discriminant Efficiency

Table 4 reports the AUC derived from an ROC analysis performed on each of the PFSF domain scores. In this analysis, we combined domain scores for both surgical and natural low libido and compared against domain scores for the control women. All seven domains of the PFSF demonstrate a high degree of predictive capacity using this indicator ($AUC \geq 0.89$). The individual ROC curve for each low libido group compared with its appropriate control group is given for the desire domain in Figure 2. Similar curves are seen when each population is considered individually, emphasizing the similarity in sexual

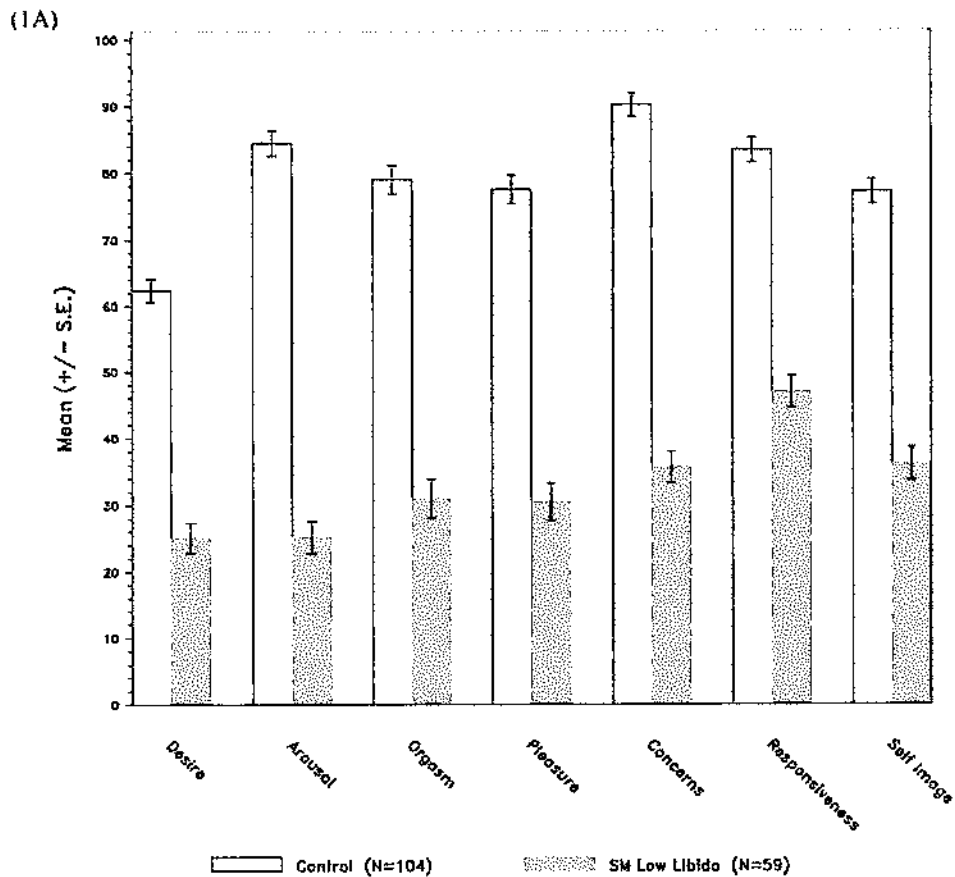


FIGURE 1. Profile of Female Sexual Function adjusted mean scores by domain. A. Low libido, surgically menopausal women ($n = 59$) and normal libido control subjects ($n = 104$); B. low libido, naturally menopausal women ($n = 88$) and normal libido naturally menopausal control subjects ($n = 47$). Error bars represent standard errors. (Continued)

(1B)

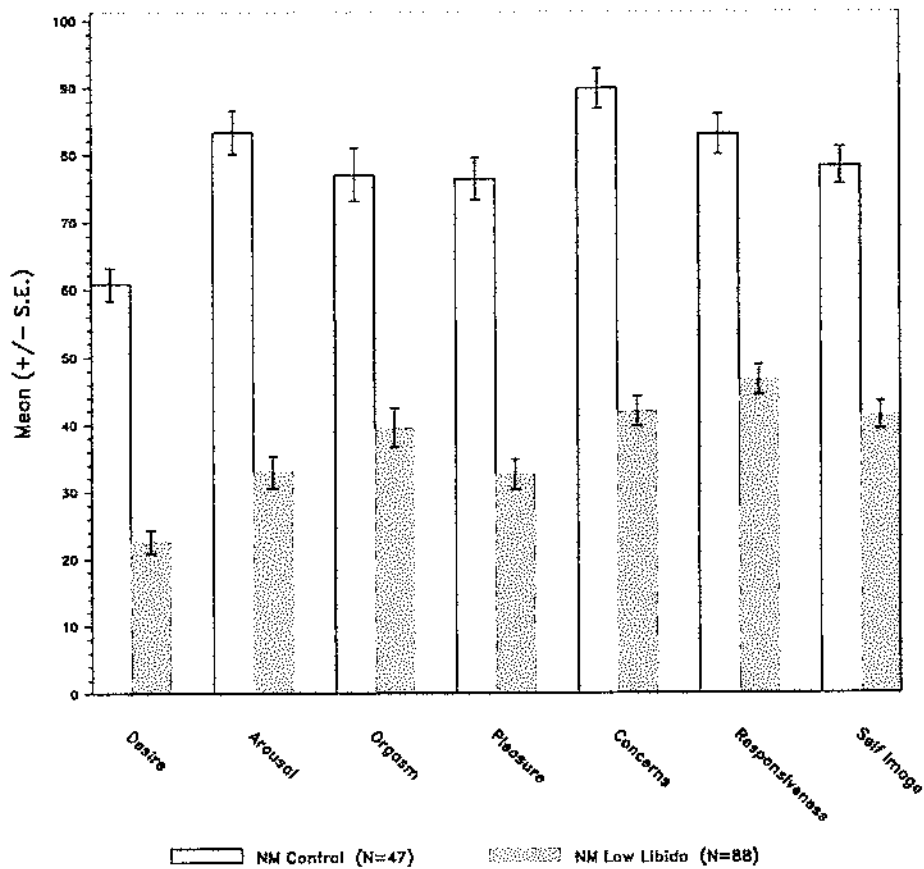


FIGURE 1. (Continued)

desire scores regardless of etiology. Similar curves were seen for the other domains (data not shown).

Table 4 also reports sensitivity and specificity estimates. We evaluated sensitivity and specificity by computing the proportion of low libido and control group women, respectively, scoring below or above a particular cut-off value for "caseness." Acceptable coefficients for sensitivity and specificity were demonstrated by all domains. The desire domain demonstrated a sensitivity of 0.94, indicating that the large majority of women positive for HSDD were identified by the scale. This predictive ability was not artifactual (that is, not constructed on the basis of a high rate of false positive errors), because the specificity was also high (0.86). In addition, the positive predictive value (PPV) for the desire score was 0.87, indicating that approximately nine out of ten predicted positives based on the PFSF desire score were true positives.

TABLE 4. ROC Analysis of Low Libido Surgically and Naturally Menopausal Women and Controls

Domain	Sensitivity	Specificity	Positive predicted value	Negative predicted value	AUC
Sexual desire	0.94	0.86	0.87	0.93	0.95
Arousal	0.93	0.92	0.92	0.93	0.97
Orgasm	0.67	0.92	0.90	0.73	0.89
Sexual pleasure	0.84	0.86	0.86	0.84	0.93
Sexual concerns	0.89	0.95	0.94	0.90	0.96
Sexual responsiveness	0.78	0.92	0.91	0.81	0.94
Sexual self-image	0.72	0.90	0.88	0.76	0.90

Convergent Validity

Table 5 shows the correlations between the domains of the PFSS and the domains of the DISF-SR. Statistically significant correlations ($p < 0.001$) were observed for all correlations of interest. The overall pattern supports the convergent validity of the PFSS. Although moderate to high correlations

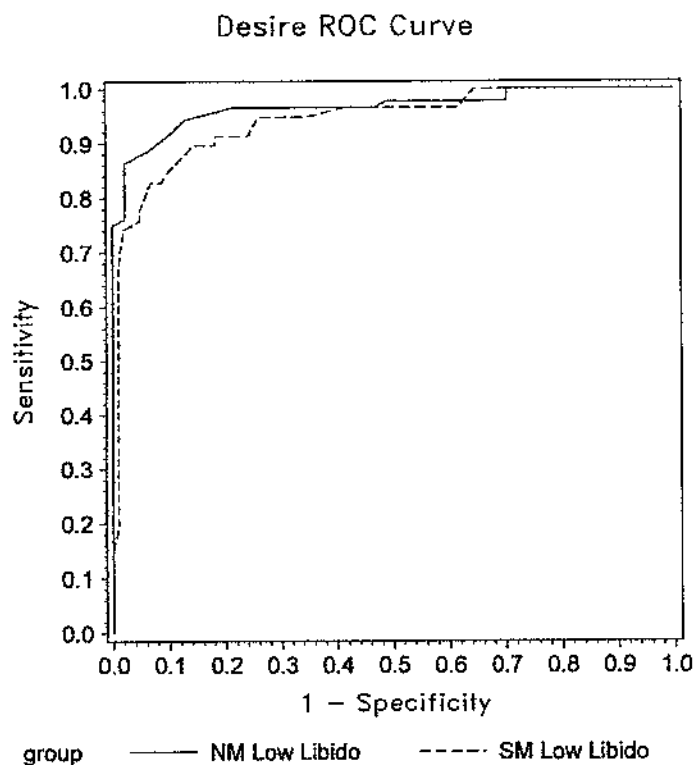
**FIGURE 2.** ROC curve for the desire domain of the Profile of Female Sexual Function. NM indicates naturally menopausal women; SM indicates surgically menopausal women.

TABLE 5. Correlation of the Profile of Female Sexual Function (PFSF) with the Derogatis Interview for Sexual Function–Self Report (DISF-SR)

PFSF domain	DISF-SR domain				
	Sexual cognition/ fantasy	Sexual arousal	Sexual behavior/ experiences	Orgasm	Drive and relationship
Sexual desire	0.64	0.76	0.67	0.69	0.82
Arousal	0.49	0.70	0.56	0.80	0.82
Orgasm	0.47	0.66	0.47	0.87	0.70
Sexual pleasure	0.51	0.73	0.56	0.83	0.81
Sexual concerns	0.45	0.67	0.55	0.74	0.78
Sexual responsiveness	0.50	0.67	0.63	0.66	0.79
Sexual self-image	0.46	0.66	0.60	0.66	0.73

were observed, no redundancy was observed, indicating that the domains measure related yet separate characteristics of sexuality.

DISCUSSION

The validation statement previously established for the PFSF, which was developed and initially validated in surgically menopausal women with low libido in North America, Europe, and Australia and in seven languages, has now been extended as a result of the current study to naturally menopausal women with low libido. The PFSF is a patient-based, self-report inventory that was developed to measure treatment benefit in menopausal women with HSDD.

The mean domain scores of both low libido groups were statistically significantly different from those of the age-matched controls for all seven domains of the instrument, demonstrating that the PFSF clearly discriminates low libido menopausal women from women free of this condition. Qualitative interviews showed that naturally menopausal women with HSDD experience the same symptoms and concerns as those expressed by surgically menopausal women with HSDD. Consistent with this observation, very similar quantitative domain scores were observed in the two populations of women, confirming the ability of the PFSF to accurately quantify these characteristics independent of the basis for menopause.

The current trial demonstrated good scale reliability for the PFSF via both test-retest and internal-consistency analyses. Slightly lower test-retest values, as compared with those seen in earlier validation studies (McHorney et al., manuscript submitted), may be a result of the somewhat long test-retest interval (4 weeks). Acceptable reliability was demonstrated independently in each low libido sample separately, as well as in the control population, and not as a function of combining heterogeneous populations, which could lead to higher test-retest reliability.

The PFSF also demonstrated admirable predictive efficiency in assigning low libido versus control subjects to the correct groups and showed overall high predictive validity in discriminating low libido from control women, as evidenced by AUC values of 0.89 or greater. Additionally, we saw moderate to high sensitivity, specificity, and positive and negative predicted values for each domain. The preliminary cutoff point used in determining these estimates was derived from previous validation studies assuming a 50% prevalence rate. Although reports of prevalence rates in the literature suggest a lower rate of HSDD, a 50% prevalence rate may be appropriate for a clinical population (Laumann et al., 1999).

In addition, good convergent validity was illustrated through correlations with the DISF, one of the standard clinical outcomes measures in the field. High correlations on a priori hypothesized construct relationships represent another strong validation statement for the PFSF.

Current longitudinal studies to assess the PFSF's ability to detect changes in sexuality due to treatment as well as to characterize domain scores in a multitude of populations are ongoing and will be reported at a later date. On the basis of this ongoing work, we determine that refinements or modifications may be made to the PFSF. To date, the instrument has been validated only in menopausal women; additional research would be required to establish the usefulness of the instrument in premenopausal women. It is also important to note that this instrument has been designed to be implemented as an outcomes measure in clinical trials to assess therapeutic efficacy. Additional studies would be required to develop a version of the PFSF for use as an assessment tool in an individual practice setting.

This study demonstrates that the PFSF is reliable and valid for assessment of sexual function in both surgical and naturally menopausal women with hypoactive sexual desire disorder.

REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, D.C.: Author.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *BDI-II manual*. San Antonio, TX: The Psychological Corporation.
- Derogatis, L. R. (1997). The Derogatis Interview for Sexual Functioning (DISF/DISF-R): An introductory report. *Journal of Sex & Marital Therapy*, *23*, 291-304.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology*, *143*, 29-36.
- Lauman, E. O., Paik, A., & Rosen, R. C. (1999). Sexual dysfunction in the United States: Prevalence and predictors. *Journal of the American Medical Association*, *281*, 537-544.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420-428.