

## Profile of Female Sexual Function: a patient-based, international, psychometric instrument for the assessment of hypoactive sexual desire in oophorectomized women

Colleen A. McHorney, PhD,<sup>1</sup> John Rust, PhD,<sup>2</sup> Susan Golombok, PhD,<sup>2</sup> Susan Davis, MD,<sup>3</sup> Celine Bouchard, MD,<sup>4</sup> Candace Brown, MD,<sup>5</sup> Rosemary Basson, MD,<sup>6</sup> C. Donati Sarti, MD,<sup>7</sup> James Kuznicki, PhD,<sup>8</sup> Cynthia Rodenberg, PhD,<sup>8</sup> and Leonard Derogatis, PhD<sup>9</sup>

### ABSTRACT

**Objective:** The purpose of this study was to develop a self-administered, patient-based questionnaire to assess loss of sexual desire and associated symptoms in postmenopausal women with hypoactive sexual desire disorder (HSDD) experiencing distress.

**Design:** Preliminary items and domains of sexual function were identified through individual and focus group interviews with postmenopausal women in the United States and Europe. A subset of items was selected for translation and further analysis. Cognitive interviews were conducted with women with HSDD and non-HSDD women in eight countries to ensure items would have the same meaning in seven languages. The resulting instrument was tested in 325 oophorectomized women with HSDD and 255 age-matched nonoophorectomized control women in the United States, Canada, Europe, and Australia.

**Results:** Psychometric item reduction analyses resulted in 37 items organized into seven domains characterizing female sexual function in postmenopausal women with HSDD. Excellent reliability and validity of the domains of the Profile of Female Sexual Function (PFSF) were observed in all geographic areas tested. Statistically significant differences between oophorectomized women with low libido and control women were found for all domains and all geographic areas.

**Conclusions:** The PFSF is a new instrument specifically designed for measurement of sexual desire in oophorectomized women with low libido. Robust psychometric properties have been established in a large number of geographic regions and languages, making it useful for assessing therapeutic change in multinational clinical trials.

**Key Words:** Hypoactive sexual desire disorder – Libido – Female sexuality – Sexual function – Psychometric instrument.

Received June 9, 2003; revised and accepted October 20, 2003.

From the <sup>1</sup>Indiana University School of Medicine and Regenstrief Institute Inc., Indianapolis, IN; <sup>2</sup>City University, London, UK; <sup>3</sup>The Jean Hailes Foundation, Clayton, Victoria, Australia; <sup>4</sup>Center Médical Santé Femme, Québec, QC, Canada; <sup>5</sup>University of Tennessee Health Science Center, Memphis, TN; <sup>6</sup>Vancouver Hospital, Vancouver, British Columbia, Canada; <sup>7</sup>Unità Operativa di Ginecologia e Ostetricia, Policlinico Monteluce, Perugia, Italy; <sup>8</sup>Procter & Gamble Pharmaceuticals Inc., Cincinnati, OH; and <sup>9</sup>University of Maryland, Baltimore, MD.

This work was supported by a grant from Procter & Gamble Pharmaceuticals Inc., Cincinnati, OH.

Address correspondence to: Cynthia Rodenberg, PhD, Procter & Gamble Pharmaceuticals, 8700 Mason-Montgomery Road, Mason, OH 45040. E-mail: rodenberg.ca@pg.com.

Sexual dysfunction is a multidimensional problem with biological, psychological, and interpersonal determinants. Community studies in which women self-report sexual dysfunction (usually via questionnaires) have found rates of sexual dysfunction in the range of 8% to 50%.<sup>1,2</sup> Focusing on dysfunction that is perceived by the woman to be a problem may reduce these rates by approximately one third.<sup>3</sup> Early attempts to characterize the female sexual response<sup>4,5</sup> are now being reexamined, resulting in new models that provide more complete descriptions of the psychological and physiological components of female

sexual function and diagnostic criteria for different types of female sexual dysfunction.<sup>6,7</sup>

General categories of female sexual dysfunction, as recognized by the Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)<sup>8</sup> include desire disorders (ie, hypoactive sexual desire disorder, HSDD), arousal disorders, orgasmic disorders, and sexual pain disorders. These disorders can occur individually or in combination and may be primary or secondary to other physical or psychiatric conditions. The DSM-IV and the recent international consensus conference on female sexual dysfunction both recognize the need to include the level of personal distress as a diagnostic criterion.<sup>6,8</sup>

Women who have undergone a surgical menopause (bilateral oophorectomy) may experience decreased levels of sexual desire after the procedure, perhaps as a result of loss of ovarian sex steroid production.<sup>9</sup> Such women commonly report feelings of loss, dissatisfaction, or distress as a result of their diminished libido<sup>10</sup> and, therefore, can be classified as having hypoactive sexual desire disorder (HSDD).

The development of clinical treatments to address HSDD in oophorectomized women requires that a suitable psychometric instrument be available to measure aspects of desire in this population.<sup>11</sup> At the time we began our work, a number of instruments existed for the measurement of female sexual function, including the Golombok-Rust Inventory of Sexual Satisfaction,<sup>12</sup> the Brief Index of Sexual Function for Women,<sup>13,14</sup> and the Derogatis Interview for Sexual Function.<sup>15</sup> However, none of these instruments had been developed and validated specifically for use in evaluation of treatment response in postmenopausal women with HSDD in multinational clinical trials.

Our objective was to develop a patient-based and psychometrically sound instrument, the Profile of Female Sexual Function (PFSF), to measure aspects of sexual desire in oophorectomized women specifically experiencing low libido. Hence, all participants were estrogen replete to exclude vasomotor instability or vaginal atrophy as primary causes of diminished sexual interest. Women, age-matched by decade, who were totally satisfied with their level of sexual interest and activity participated as healthy comparators. Because of its intended use in clinical trials in women with HSDD, the instrument was constructed to reflect the medical diagnosis of this disorder as well as its effect on women's thoughts, feelings, emotions, and behaviors. Further, the instrument was to be developed and validated in multiple geographic regions and languages so as to be useful in multinational clinical trials. Initial versions

of this instrument were developed by selecting inventory items from among a large pool of potential items generated by women via focus groups. After translation, linguistic validation, and international harmonization, the instrument was evaluated in studies in the United States, Canada, Europe, and Australia to further refine its content and determine its psychometric properties and potential usefulness in multiple geographic areas. We report the development of the PFSF herein.

## METHODS

### Initial item development and linguistic validation of the PFSF

The preliminary items and domains of sexual function important to oophorectomized women with low libido were identified via a series of individual interviews and focus group interviews with 91 surgically and 25 naturally postmenopausal women with HSDD in four countries in Europe (Germany, UK, Italy, and France) and in the United States. Eighty-nine individual interviews were conducted, and 27 additional women participated in two focus groups. In addition, individual interviews were conducted with 96 physicians in the United States and England to determine if the women's needs matched physicians' perceptions of their needs. Approximately 400 items generated from the interviews were reviewed for redundancy, language level, freedom from slang or jargon, apparent meaning, and relevance to women. A subset of 83 items, designed to ensure that essential elements of sexuality and related distress were addressed, was selected for further evaluation. Of the 83 items, 17 relating to aspects of distress were evaluated as a separate instrument, and 66 formed the basis for the development of the PFSF. This item development process was conducted by a group of clinicians and psychometricians with experience in the treatment or measurement of sexual dysfunction. This group included individuals with expertise in clinical psychology, statistics, psychometrics, public health, and consumer research.

### Content validity: qualitative linguistic validation and international harmonization

Cognitive interviews were used to assess linguistic validity across multiple geographic areas and languages. After translation of items from English, we conducted interviews with 55 surgically postmenopausal women with HSDD and six nonoophorectomized women without HSDD in the United States, Germany, Holland, France, Italy, Canada, Australia, and the UK, ages ranging from 34 to 70 years, to ensure that

each item would have the same meaning to women across the languages of interest: English as spoken in the United States, England, and Australia, native and Canadian French, Dutch, German, Italian, and Spanish as spoken in the United States. Professional interviewers and linguists conducted these interviews of native speakers of each language before quantitative evaluation of the inventories. After the interviews, a group of linguists reviewed the items across languages to ensure consistent meanings in all languages.

Cognitive laboratory methods use a variety of techniques, in individual or group discussion formats, to identify the process by which people process, interpret, and answer questionnaire items.<sup>16-19</sup> Cognitive interviews of all 61 women were conducted to determine the women's understanding of the items and to identify items with ambiguous or multiple meanings. The interviews were also used to determine the exact wording in each language required to preserve the meaning of the items and to establish that the item content was relevant in each region. In this instance, women were asked to "think aloud" as they answered the questionnaire items. Examples of multiple and ambiguous meanings identified during this process are shown in Table 1.

#### Inventory evaluation

After translations and the cognitive interviews, the instrument was further evaluated in three nonrandomized, parallel-group, multicenter validation studies conducted in the United States, Canada, Europe, and Australia. These trials enrolled surgically postmenopausal women with low libido and age-matched (within geographical region) control women who had intact ovaries and normal libido. Inclusion and exclusion criteria consistent with the DSM-IV criteria for hypoactive sexual desire disorder<sup>8</sup> were used to determine low-libido status. These women had to report a satisfying sex life before menopause (either surgically or

naturally) with a meaningful loss in desire and decrease in sexual activity after menopause, and they had to be experiencing concern over their decreased level of desire for sexual activity. The control group consisted of women who characterized their sex life as good and satisfying, were satisfied with their current level of sexual activity, and were not concerned with their level of interest in sex. Age matching was performed by decade ranges. The ratio of low-libido women to controls was 2 to 1 in the US study and 1 to 1 in the other two studies.

All women were between 20 and 70 years of age and were required to be in a stable, monogamous relationship for at least 1 year before study entry with a partner who was sexually functional. Women in the surgically postmenopausal, low-libido group were to have undergone bilateral salpingo-oophorectomy and total hysterectomy at least 1 year before study entry, whereas women in the control group were to have at least one ovary. All postmenopausal women, regardless of group status, were on stable estrogen therapy alone or estrogen with progestin therapy for at least 3 months before study entry.

Women were excluded if they had chronic or acute life stress relating to major life change that could have interfered with sexual activity. Women whose disorder might be accounted for by another Axis I disorder (DSM-IV criteria)<sup>8</sup> or by use of a substance or other medical condition were excluded. Thus, women who had depression (Beck Depression Inventory-II score  $\geq 14$ ),<sup>20</sup> dyspareunia, physical problems or previous sexual trauma that might interfere with sexual activity, or were taking medications known to affect desire (eg, selective serotonin reuptake inhibitors, tricyclic antidepressants, antiandrogens, beta-blockers, dehydroepiandrosterone) were excluded.

The duration of the study for each woman was 4 weeks from the date of her study enrollment. Women visited the study centers at baseline and after 2 and 4 weeks to complete study procedures. At baseline, each

TABLE 1. Examples of multiple or ambiguous meanings identified in participant interviews during the translation and linguistic validation process

Item/domain	Intended meaning	Unintended meaning/result
"I got wet during sex"	Intended to measure vaginal lubrication, an aspect of arousal	Some women interpreted "wetness" as referring to urination during sex, perspiration during sex, or getting wet with ejaculate
"I feel like a sexual person"	Intended to reflect sexual self image	Was interpreted after translation as "I feel like a prostitute"
"I feel relaxed about sex"	Intended to measure comfort with idea of having sex	Some women reported feeling relaxed about sex, but were not having sex
Spontaneity domain (items relating to sexual activity with little or no planning)	Expected to be higher in women with normal libido, compared with women with low libido	Women with low libido reported high levels of spontaneity; unlike women with normal libido, they rarely thought about or planned sex, which was usually initiated by the partner and viewed as spontaneous

woman completed screening procedures, provided demographic data, and read through the inventories to be administered later in the study. The PFSF was completed at the 2- and 4-week visits.

#### Item reduction and domain determination analyses

Initial item reduction was performed using the data from the US study to determine the most effective subset of items for inclusion in the final instrument and to identify the underlying factors characterizing sexual functioning in the participant population. Standard psychometric item analyses were performed using the response from the first administration of the PFSF.<sup>21-24</sup> Data from the Canadian and European/Australian studies were then analyzed to determine whether the US results could be applied to these regions and to identify any modifications that needed to be made.

Reducing the length of the PFSF involved making a trade-off between breadth and depth of coverage of aspects of female sexual dysfunction. Because we did not want to sacrifice either breadth of measurement (comprehensiveness) or depth of measurement (precision), we decided to assess numerous domains using as few items as possible to construct each multi-item scale.

We used a variety of well-accepted item analysis techniques to reduce the length of the PFSF.<sup>21</sup> We assessed unidimensionality by extracting principal components from the correlations among the items. The principal components analysis was used to identify items measuring a single general factor. Items with correlations of less than 0.40 with the first or second unrotated principal component were considered for removal. The first two components were examined to account for possible differential loading of positively- and negatively-worded items. In addition, a rotated factor analysis, using a varimax rotation, was used to identify factor loadings for each item. Analyses were performed on the raw data and on a Z-score transformation to account for any possible acquiescent responding by participants. Items showing complex correlations across multiple factors and items having no meaningful correlations on any factor were considered for removal. Finally, for each domain, redundant items within a domain (inter-item correlation > 0.75) and items not correlating significantly with their domain total scores (item-total correlation corrected for overlap < 0.40) were also considered for removal.

Items with missing data rates greater than 10% were candidates for exclusion. Item frequency distributions were calculated, and the range of responses for each item was assessed. Item distributions were examined for floor and ceiling effects, ie, the respective clustering

of participants at the worst- and best-possible response categories. Ceiling effects are undesirable because participants cannot improve their scores if they are at the ceiling. A floor effect in the control group indicates that even participants in the control group did not identify with that item. Items that had more than 25% of responses in either the bottom or top two categories were candidates for exclusion.

Items that did not discriminate well between the low-libido and control groups were identified for possible deletion. To account for any variation between the two groups, an analysis of covariance with region, source, marital status, length of relationship with partner, and age covariate adjustment was used.

#### Scale Construction

Each item of the PFSF, except the global "satisfaction with sexuality" question, was rated by participants on a six-point categorical rating scale, with "Always" corresponding to a score of 1, "Very Often" to 2, "Often" to 3, "Sometimes" to 4, "Seldom" to 5, and "Never" to 6. The global "satisfaction with sexuality" question—which asked: "Considering the past 30 days, how would you rate your overall satisfaction with your sexuality?"—had a five-point category rating scale, with "Poor" corresponding to a score of 1, "Fair" to 2, "Good" to 3, "Very Good" to 4, and "Excellent" to 5. Each scale (Sexual Desire, Arousal, Orgasm, Sexual Pleasure, Sexual Concerns, Sexual Responsiveness, and Sexual Self Image) was constructed using Likert's method of summated ratings, which equally weighs each item and sums them into an overall scale score. All scales were linearly transformed to a 0 to 100 metric, with 100 indicating the most favorable state (best sexual function), 0 the least favorable (worse sexual function), and scores in between representing the percentage of the total possible score achieved.

#### Reliability and validity analyses

After completion of the item reduction phase and identification of final domains, the psychometric properties of reliability and validity were assessed for each PFSF domain and for the overall satisfaction with sexuality question. An intraclass correlation coefficient and Cronbach's alpha value for each scale score was computed to evaluate test-retest reliability and internal-consistency reliability, respectively. The intraclass correlation coefficient was obtained from the variance components of a repeated-measures analysis of variance model.<sup>25</sup>

Classical known-groups validity<sup>23</sup> was assessed by evaluating the ability of the instrument to discriminate

between hysterectomized, oophorectomized women with low libido and nonoophorectomized control women. A repeated-measures analysis, adjusted by site, age, source, marital status, and length of relationship, was used to test the hypothesis of no difference between group means. Correlations among all domains of the PFSF were determined using Pearson correlation coefficients to evaluate the extent to which the domains measure unidimensional aspects of sexuality.

## RESULTS

### Characteristics of participants in the quantitative studies

Demographic and baseline characteristics of the women enrolled in the three quantitative studies conducted in the United States (18 centers), Canada (6 centers), and Europe/Australia (14 centers) are shown in Table 2. In each study, women in the low-libido and control groups were similar in race, height, and weight. Low-libido women were slightly older, slightly more likely to be married to their partners, and in relationships of slightly longer duration than those in the control group. The mean length of relationship was greater than a decade in all studies. Because all women in the low-libido group had undergone hysterectomy and bilateral oophorectomy, they were all postmenopausal; 28% to 35% of the women in the control group were postmenopausal. Participants in the European/Australian study were slightly more likely to be married and in relationships of longer duration than participants in the other studies.

### Item reduction

Initial item reduction was undertaken using the data from the US study alone. Similar results were obtained

when data from the Canadian and European/Australian studies were examined.

### Data quality and item distributions

Data quality was assessed by examining missing data rates and floor and ceiling effects. Overall, missing data rates were low (< 2.3% in the US study), and no items were removed from the inventory on that basis. Ten items showed evidence of floor effects in the control group or ceiling effects in the low-libido group or had low discriminability and were deleted.

### Domain identification

Selection of domains and items in domains was based on principal components and factor analysis, item-domain correlations, inter-item correlations, and clinical considerations of female sexual dysfunction in general and hypoactive sexual desire disorder in particular.

Principal components analysis was performed to identify any item that did not seem to be related to a general sexuality factor. We examined the results of an unrotated principal components analysis and identified items that did not load on either of the first two components. One item from the PFSF was deleted for not having a meaningful loading (loading of 0.12 and 0.11, respectively) on either of the first two principal components.

Varimax rotated factor analysis of the raw data identified seven domains that were labeled Sexual Pleasure, Sexual Desire, Responsiveness, Arousal/Orgasm, Sexual Self Image, Sexual Concerns, and Disinterest. Thirteen items were deleted from the PFSF because of low or multiple loadings. The Disinterest domain contained two items and accounted for less than 2.5% of total variance in the US data and was, therefore, deleted.

TABLE 2. Characteristics of the low libido women and controls in the three quantitative development studies

	US study		Canadian study		European/Australian study	
	Low-libido group (n = 224)	Control group (n = 146)	Low-libido group (n = 35)	Control group (n = 33)	Low-libido group (n = 224)	Control group (n = 146)
Age in years [mean (SE)]	48 (0.55)	44 (0.76)	48 (1.47)	45 (1.65)	50 (0.79)	47 (0.86)
Height in cm [mean (SE)]	164.5 (0.5)	164.1 (0.5)	162.4 (1.0)	161.9 (0.9)	162.8 (0.6)	163.9 (0.7)
Weight in kg [mean (SE)]	67.7 (0.6)	65.5 (0.8)	64.8 (1.7)	62.5 (1.7)	66.1 (1.1)	64.7 (1.1)
Race [n (%)]						
Caucasian	198 (88)	128 (88)	35 (100)	31 (94)	88 (98)	80 (98)
Black	21 (9)	14 (10)	0	0	0	0
Other	5 (2)	4 (3)	0	2 (6)	2 (2)	2 (2)
Married [n (%)]	181 (81)	105 (72)	24 (69)	18 (55)	79 (88)	70 (85)
Mean length of relationship in years [mean (SE)]	16.7 (0.77)	13.6 (0.91)	14.2 (1.84)	11.9 (1.89)	22.8 (1.21)	20.7 (1.17)
Women who were postmenopausal [n (%)]	224 (100)	41 (28)	35 (100)	11 (33)	90 (100)	29 (35)

PROFILE OF FEMALE SEXUAL FUNCTION

Table 3 shows the results from the factor analysis for those items not deleted.

Further considerations led to the modification of this structure. Arousal and orgasm are medically recognized as separate aspects of sexuality (DSM-IV).<sup>8</sup> For this reason, despite the documented comorbidity of arousal and orgasm dysfunction,<sup>26</sup> Arousal and Orgasm were treated as separate domains. Items 14, 56, 60, and 65 loaded on two or more dimensions but were retained based on clinical considerations. The remaining items were arranged into the PFSF with the following domains (number of items): Desire (9), Arousal (3), Orgasm (4), Pleasure (9), Sexual Concerns (3), Responsiveness (7), and Sexual Self Image (4). Items belonging to the particular domains are indicated in bold in Table 3.

Data quality and domain identification analyses using data from the Canadian and European/Australian studies produced similar results (data not shown). Inter-item correlations computed among items within domains indicated redundancy among items in the Pleasure domain for all three studies, and two additional items (items 49 and 51) were, therefore, deleted from that domain. The final instrument contained 37 items in seven domains (Table 4), and the additional question on global "satisfaction with sexuality."

Score distributions

The score distributions for the final form of the instrument using low-libido participants from all studies are given in Table 4. Overall, the domains of the PFSF

TABLE 3. Factor analysis results for preliminary items and domains evaluated during PFSF development

Item	Abbreviated item content	Pleasure	Desire	Responsiveness	Arousal/orgasm	Self-image	Concerns	Disinterest
2	I felt like having sex (R)	---	<b>0.74</b>	--	-	-	-	-
5	My sexual desire was high (R)	-	<b>0.76</b>	-	-	-	-	-
6	I really wanted sex (R)	-	<b>0.76</b>	-	-	-	-	-
9	I felt sexual desire (R)	-	<b>0.73</b>	-	-	-	-	-
11	I lacked sexual desire	-	<b>0.58</b>	-	-	-	-	-
13	I had strong sexual feelings (R)	-	<b>0.73</b>	-	-	-	-	-
14	I was uninterested in sex	-	<b>0.52</b>	<b>0.50</b>	-	-	-	<b>0.44</b>
17	I got warm all over just thinking about sex (R)	-	<b>0.65</b>	--	-	-	-	-
18	It was very difficult for me to become aroused	-	<b>0.46</b>	--	<b>0.66</b>	-	-	-
19	I felt really excited sexually (R)	-	<b>0.68</b>	-	-	-	-	-
20	I really had to focus to get sexually aroused	-	-	-	<b>0.61</b>	-	-	-
22	Getting aroused took forever	-	-	-	<b>0.65</b>	-	-	-
26	I initiated sex (R)	-	-	<b>0.55</b>	-	-	-	-
28	I made up excuses to avoid having sex	-	-	<b>0.73</b>	-	-	-	-
31	I avoided doing anything that would get my partner sexually excited	-	-	<b>0.78</b>	-	-	-	-
33	I welcomed the chance to have sex (R)	-	-	<b>0.57</b>	-	-	-	-
34	I avoided having sex	-	-	<b>0.79</b>	-	-	-	-
37	I got a lot of pleasure from sex (R)	<b>0.80</b>	-	-	-	-	-	-
38	Sex felt good (R)	<b>0.79</b>	-	-	-	-	-	-
39	I was frustrated about my sex life	-	-	-	-	-	<b>0.79</b>	-
40	I felt sexually numb	<b>0.50</b>	-	-	-	-	<b>0.48</b>	-
41	I felt warm all over during sex (R)	<b>0.74</b>	-	-	-	-	-	-
43	Sex was wonderful (R)	<b>0.84</b>	-	-	-	-	-	-
44	I felt distressed about sex	-	-	-	-	-	<b>0.69</b>	-
45	I felt tingly all over during sex (R)	<b>0.73</b>	-	-	-	-	-	-
49	Sex was fulfilling (R)	<b>0.83</b>	-	-	-	-	-	-
50	Sex was exciting (R)	<b>0.72</b>	-	-	-	-	-	-
51	Sex gave me great feelings all over my body (R)	<b>0.82</b>	-	-	-	-	-	-
52	I dreaded having sex	-	-	<b>0.77</b>	-	-	-	-
53	Sex was satisfying (R)	<b>0.81</b>	-	-	-	-	-	-
55	Having sex was a chore	-	-	<b>0.74</b>	-	-	-	-
56	I reached orgasm easily (R)	<b>0.64</b>	-	-	<b>0.48</b>	-	-	-
57	It took a lot of work for me to reach orgasm	-	-	-	<b>0.84</b>	-	-	-
59	Having orgasms was difficult	-	-	-	<b>0.83</b>	-	-	-
60	Reaching orgasm was impossible	<b>0.44</b>	-	-	<b>0.62</b>	-	-	-
61	Good about yourself sexually (R)	-	-	-	-	<b>0.85</b>	-	-
62	Sexually desirable (R)	-	-	-	-	<b>0.82</b>	-	-
63	Like a sensuous woman (R)	-	-	-	-	<b>0.85</b>	-	-
65	Unhappy about yourself sexually	-	-	-	-	<b>0.44</b>	<b>0.54</b>	-

Data is from the US study. R, indicates that scoring was reversed before analysis; -, represents factor loading < 0.4. Items in bold have been grouped into a domain based on statistical and clinical considerations. PFSF, Profile of Female Sexual Function.

**TABLE 4.** Features of score distributions in low-libido population from all three studies (n = 325) and sample item content

PFSF domain	Items (n)	Levels (n)	Mean	SD	Median	% floor	% ceiling	Sample item
Desire	9	45	26	16	24	4	0	"I had strong sexual feelings"
Pleasure	7	35	37	24	34	6	2	"I got a lot of pleasure from sex"
Responsiveness	7	35	47	24	49	2	0	"I made up excuses to avoid having sex"
Arousal	3	15	34	26	33	14	1	"It was difficult for me to become aroused"
Orgasm	4	20	40	26	40	9	1	"I reached orgasm easily"
Sexual self-image	4	20	42	22	40	2	< 1	"I felt sexually desirable"
Sexual concerns	3	15	45	27	47	8	4	"I was frustrated about my sex life"
Global "satisfaction with sexuality" question	1	5	1.9	0.9	2	39	< 1	"How would you rate your overall satisfaction with your sexuality?"

PFSF domain scores range from 0 (worst sexual function) to 100 (best sexual function). Global "satisfaction with sexuality" scores range from 1 (poor) to 5 (excellent). PFSF, Profile of Female Sexual Function.

**TABLE 5.** Reliability estimates for the Profile of Female Sexual Function

Low-libido group domain	Internal consistency and test-retest reliability					
	US study (n = 208)		Canadian study (n = 33)		European/Australian study (n = 84)	
	Cronbach's $\alpha$	ICC	Cronbach's $\alpha$	ICC	Cronbach's $\alpha$	ICC
Desire	0.92	0.83	0.92	0.81	0.92	0.83
Pleasure	0.94	0.83	0.96	0.72	0.96	0.87
Responsiveness	0.91	0.88	0.86	0.90	0.89	0.86
Arousal	0.89	0.78	0.83	0.69	0.95	0.83
Orgasm	0.84	0.84	0.90	0.52	0.93	0.80
Sexual self-image	0.83	0.71	0.81	0.78	0.88	0.85
Sexual concerns	0.79	0.77	0.85	0.77	0.84	0.85

ICC, intra-class correlation coefficient.

are symmetrically distributed with mean values similar to the median values. All the domains except Arousal and the global "satisfaction with sexuality" question had inconsequential floor effects, indicating the ability of the PFSF to measure gradients in disease status. Additionally, all domains had no more than 4% of the participants experiencing a ceiling effect, indicating the ability of the PFSF to measure positive changes in sexuality.

#### Scale reliability

Reliability estimates in the low-libido group are presented in Table 5. In low-libido women, the alpha coefficients ranged from 0.79 to 0.94 in the United States, 0.81 to 0.96 in Canada, and 0.88 to 0.96 in Europe/Australia. A similar range (0.5-0.96) of values was observed in the control group, with all but two of 24 values equal to or greater than 0.8. Two-week retest reliability estimates were all substantial to excellent. Intraclass correlation coefficients (ICCs) were acceptable in all regions and indicated consistency of response across the two administrations. A similar range (0.66-0.92) of ICC values was observed in the control group, with all but one of 24 values equal to or greater than 0.73. Slightly lower but acceptable ICC values

were seen for the global "satisfaction with sexuality" question (ICC range from 0.48-0.79). This value is typical for single-item measures.

#### Domain-domain correlations

Table 6 presents scale-scale correlations in the low-libido group in the United States. Correlations ranged from a low of 0.18 to a high of 0.66 (median, 0.50). Hence, no redundancy was observed, indicating that the domains measure related yet separate characteristics of sexuality. The lowest correlations were observed between Orgasm and Responsiveness and between Orgasm and Sexual Self Image. The highest correlations were observed between Desire and Responsiveness and Desire and Arousal. A similar range of values was observed in the control group (data not shown). A similar range of values was seen in both the Canadian and European/Australian studies (data not shown).

#### Clinical known-groups validity

The PFSF was evaluated for its ability to discriminate between low-libido and control participants. Figure 1 shows the mean (SE) scores for all domains of the PFSF from the first administration for low-libido and

PROFILE OF FEMALE SEXUAL FUNCTION

TABLE 6. Domain intercorrelations of the Profile of Female Sexual Function

Low-libido group (n = 208)	Desire	Pleasure	Responsiveness	Arousal	Orgasm	Sexual self-image	Sexual concerns
Desire	1.00						
Pleasure	0.54	1.00					
Responsiveness	0.56	0.46	1.00				
Arousal	0.64	0.55	0.51	1.00			
Orgasm	0.29	0.54	0.18	0.58	1.00		
Sexual self-image	0.41	0.39	0.49	0.29	0.18	1.00	
Sexual concerns	0.45	0.55	0.53	0.54	0.40	0.48	1.00

All data from US study; based on the first administration of the questionnaire.

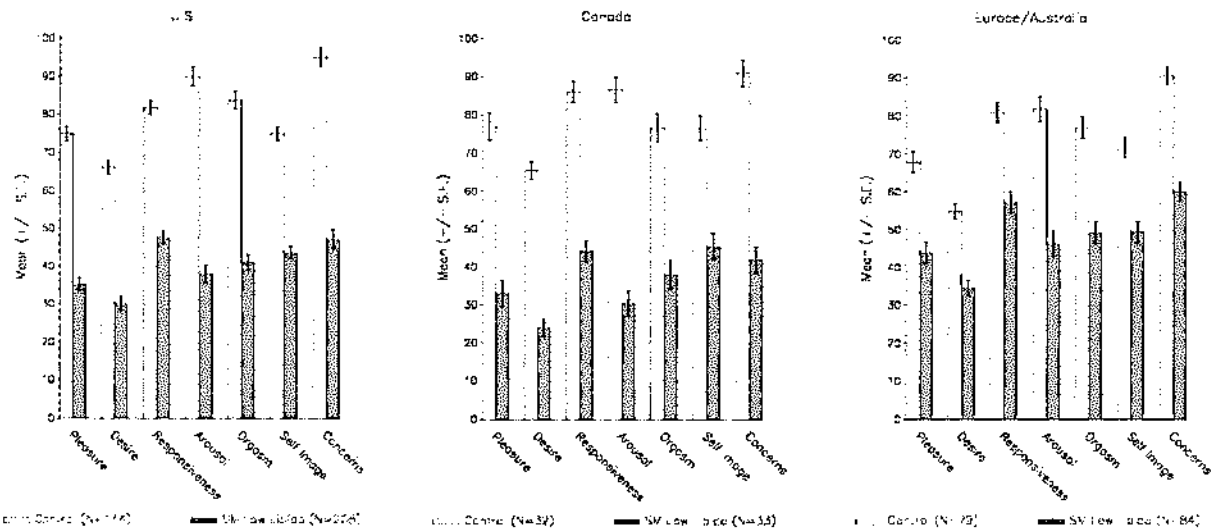


FIG. 1. Discriminant ability of the PFSF by domain across geographic regions. Statistically significantly lower scores were seen in the low-libido group compared with the control group for all geographic regions and all PFSF domains.

control group participants from the US, Canadian, and European/Australian studies. For all regions, women with HSDD scored significantly lower than the control women ( $P < 0.001$ ). Additionally, a similar profile of response was observed over the geographic regions supporting an expectation from qualitative content validation that the same content would be captured regardless of cultural differences. The instrument also showed good discriminability in younger and older women (data not shown). Additionally, for each study, women in the surgically postmenopause low-libido groups reported significantly ( $P < 0.0001$ ) less general satisfaction with their sexuality than their respective control groups (data not shown).

DISCUSSION

The Profile of Female Sexual Function is a patient-based, multinational, psychometrically validated instrument for the measurement of loss of sexual desire

and related symptoms in oophorectomized women with low libido. Unlike already existing self-report measures, development of the PFSF entailed extensive qualitative content and multinational linguistic validation and extensive multinational quantitative validation in a population of postmenopausal women with HSDD.

Analysis of data from quantitative tests indicates that the PFSF demonstrates robust psychometric properties across numerous geographic regions, including good inter-item correlations, discriminant ability, test/retest reliability, internal-consistency reliability, and construct validity (including factorial and know-groups validity). We consider this instrument to be an appropriate tool for the assessment of response to treatment in clinical trials of oophorectomized women with HSDD.

A central element of this instrument development program was the use of women's input to ensure that the instrument would be relevant to women and would accurately describe their symptoms, feelings, behaviors, and attitudes. The numerous ambiguities, multiple



meanings, and unintended meanings discovered and corrected during the process of developing the PFSF demonstrate the importance of careful attention to content validity.

Translation and adaptation of instruments across cultures requires special efforts to ensure that the test content and meaning remains the same.<sup>27</sup> The use of cognitive interviews in multiple languages and cultures was a critical element in the development of this instrument. Expressions of symptoms, feelings, behaviors, and attitudes can and do mean different things to different women, and the meaning can differ from what was intended, especially after translation. Items that had unclear or multiple meanings were identified and either eliminated, modified by participant input, or revised based on quantitative analyses. This attention to ensuring content validity greatly enhances confidence that the PFSF measures what it is intended to measure in all the languages and cultures in which it was developed.

The psychological experience of sexuality in postmenopausal women with low libido, as measured by this instrument, was remarkably similar in the three regions studied. It is not surprising that women with low libido differed from women in the control group on all domains of the PFSF, including those not directly measuring sexual desire, because hypoactive sexual desire is known to affect many aspects of sexuality.<sup>26, 28-30</sup> In addition, generation of items by the target population ensured that only those symptoms, feelings, behaviors, and attitudes experienced in each domain by women with hypoactive sexual desire disorder are included in the instrument. The similarity of response across regions supports the generalizability of the instrument's content in all the geographic areas tested and provides a psychological profile of the impact of hypoactive sexual desire disorder on each domain of sexuality that it affects.

The components of hypoactive sexual desire measured in the PFSF reflect the pervasive effects of low libido on many aspects of sexual experience, including arousal, orgasm, sexual pleasure, and the distress associated with the condition as measured by the Sexual Concerns domain, with a high level of comorbidity across domains. A unique aspect of the PFSF is the measurement of Sexual Self Image and Responsiveness. The Sexual Self Image domain measures a woman's feeling about herself as a sexual person, and the Responsiveness domain measures the tendency to approach or avoid opportunities for sexual activity. Both of these aspects of low desire have been noted in the literature,<sup>26</sup> and a seemingly related concept of "Sexual

Self Esteem" has been identified and measured in college students.<sup>31</sup> However, their formal measurement by a psychometrically-sound instrument among surgically postmenopausal women seems unique to the PFSF. Their emergence in this instrument may be a manifestation of the importance of these feelings to the women who generated the instrument's content. The domains of the PFSF differ substantially from those in the BISF-W, an instrument previously developed and used to measure treatment response in women with low sexual desire.<sup>32,33</sup> For example, the Brief Index of Sexual Functioning for Women (BISF-W) dimensions of "relationship satisfaction" and "frequency" and the PFSF domain of Sexual Self Image do not have obvious corresponding dimensions in the other instrument. A validation study comparing the PFSF with the Derogatis Interviewer of Sexual Function-Self Report (DISF-SR)<sup>15</sup> has been conducted and is being reported separately.<sup>34</sup>

Results of a randomized, double-blind, placebo-controlled trial in which the PFSF was used to measure treatment response have been recently reported.<sup>35</sup> In that study, significant increases in desire, as measured by the desire domain of the PFSF, were accompanied by significant increases in the frequency of sexual activity in women with HSDD who received transdermal testosterone.

One limitation of this work to date is that the validation program of the PFSF has so far been confined to surgically postmenopausal women with HSDD. Further studies are underway to extend the validity of the PFSF in its final form to both surgically and naturally postmenopausal women with low libido. Additionally, this instrument has been designed to be used as an outcome measure in clinical studies to assess efficacy due to treatment effect and not as a clinical diagnostic tool to be used in a physician setting. Further analyses will be required to determine what modifications may be needed for the PFSF to be a useful diagnostic tool. Given that the current instrument is lengthy, further work to develop a screening tool might include attempts to create a shorter form suitable for use in clinical practice.

## CONCLUSIONS

The PFSF is a new psychometric instrument for the measurement of loss of sexual desire and related aspects of sexual function in oophorectomized women with HSDD. This instrument is based on extensive input from women and its content and statistical properties have been validated in multiple geographic regions.

**Acknowledgments:** The authors would like to thank Gloria Yiu for performing statistical analyses and Lisa Bosch for assistance in the preparation of the manuscript.

REFERENCES

1. Béjin A. Sexual pleasures, dysfunctions, fantasies, and satisfaction. In: Spira, A., Bajos, N, eds. *Sexual Behaviour and AIDS*. Aldershot, England: Avebury. 1994:163-171.
2. Kontula O, Haavio-Mannila E. *Sexual Pleasures. Enhancement of Sex Life in Finland, 1971-1992*. Aldershot, England: Dartmouth, 1995.
3. Fugl-Meyer AR, Sjögern Fugl-Meyer F. Sexual disabilities, problems, and satisfaction in 18- 74-year-old Swedes. *Scandinavian J Sexology* 1999;2:79-105.
4. Masters WH, Johnson VE. *Human Sexual Response*. Boston: Little Brown, 1966.
5. Kaplan HS. Hypoactive sexual desire. *J Sex Marital Ther* 1979;3: 3-9.
6. Basson R, Berman J, Burnett A, et al. Report of the international consensus development conference on female sexual dysfunction: definitions and classifications. *J Urol* 2000;163:888-893.
7. Basson R. Rethinking low sexual desire in women. *Br J Obstet Gynecol* 2002;109:357-363.
8. American Psychiatric Association. *DSM-IV: Diagnostic and statistical manual of mental disorders*, 4<sup>th</sup> ed. Washington, DC: American Psychiatric Association, 1994.
9. Nathorst-Boos J, von Schoultz B. Psychological reactions and sexual life after hysterectomy with and without oophorectomy. *Gyn Obstet Invest* 1992;34:97-101.
10. Zussman L, Zussman S, Sunley R, Bjornson E. Sexual response after hysterectomy-oophorectomy: recent studies and reconsideration of psychogenesis. *Am J Obstet Gynecol* 1981;140:725-729.
11. U.S. Food & Drug Administration. Guidance for industry. Female sexual dysfunction: clinical development of drug products for treatment [Draft guidance]. Available at: <http://www.fda.gov/cder/guidance/3312dfl.htm>. Accessed June 2, 2004.
12. Rust J, Golombok S. The GRJSS: A psychometric instrument for the assessment of sexual dysfunction. *Arch Sex Behav* 1986;15:153-61.
13. Taylor JF, Rosen RC, Leiblum SR. Self-report assessment of female sexual function: psychometric evaluation of the Brief Index of Sexual Functioning for Women. *Arch Sex Behav* 1994;23:627-643.
14. Rosen RC, Taylor JF, Leiblum SR, Bachmann GA. Prevalence of sexual dysfunction in women: results of a survey study of 329 women in an outpatient gynecological clinic. *J Sex Marital Ther* 1993;19:171-188.
15. Derogatis LR. The Derogatis interview for sexual functioning (DISF/DISF-SR): An introductory report. *J Sex Marital Ther* 1997; 23:291-304.
16. Jobe JB, Mingay DJ. Cognitive laboratory approach to designing questionnaires for surveys of the elderly. *Public Health Rep* 1990; 105:518-524.
17. Fienberg S, Loftus E, Tanur J. Cognitive aspects of health survey methodology: An overview. *Milbank Mem Fund Q* 1985;63:598-614.
18. Schwarz N, Sudman S. *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass, 1996.
19. Sudman S, Bradburn NM, Schwartz N. *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass, 1996.
20. Beck AT, Steer RA, Brown GK. *Manual for the Beck Depression Inventory-II*. San Antonio: The Psychological Corporation, 1996.
21. Nunnally JC, Bernstein IH. *Psychometric Theory*, 3<sup>rd</sup> ed. New York: McGraw-Hill, 1994.
22. McHorney CA, Ware JE, Rogers W, Raczek AE, Lu JFR. The validity and relative precision of MOS short- and long-form health scales and Dartmouth COOP Charts: Results from the Medical Outcomes Study. *Med Care* 1992;30:MS253-MS265.
23. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Status Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;31:247-263.
24. McHorney CA, Ware JE, Lu JFR, Sherburne CD. The MOS 36-Item Short Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.
25. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86:420-428.
26. Andersen BL, Cyranowski JM. Women's sexuality: behaviors, responses, and individual differences. *J Consult Clin Psychol* 1995; 6:891-906.
27. Geisinger KF. Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychol Assess* 1994;6:304-312.
28. Leiblum S. Definition and classification of female sexual disorders. *Int J of Imp Res* 1998;10:S104-S106.
29. Beck JG, Bozman GW, Qualtrough T. The experience of sexual desire: psychological correlates in a college sample. *J Sex Res* 1991; 28:443-456.
30. Leif HI. Evaluation of inhibited sexual desire: relationship aspect. In: Kaplan HS, ed. *Comprehensive Evaluation of Disorders of Sexual Desire*. Washington, DC: American Psychiatric Press Inc; 1985:59-76.
31. Snell WE, Papini DR. The Sexuality Scale: An instrument to measure sexual self-esteem, sexual-depression, and sexual-preoccupation. *J Sex Res* 1998;26:256-263.
32. Mazer NA, Leiblum SR, Rosen RC. The Brief Index of Sexual Functioning for Women (BISF-W): a new scoring algorithm and comparison of normative and surgically menopausal populations. *Menopause* 2000;7:350-363.
33. Shifren JL, Braunstein GD, Simon JA, et al. Transdermal testosterone treatment in women with impaired sexual function after oophorectomy. *N Engl J Med* 2000;343:682-688.
34. Derogatis L, Rust J, Golombok S, et al. Validation of the Profile of Female Sexual Function (PFSF) in surgically and naturally menopausal women. *J Sex Marital Therapy* 2004;30:26-36.
35. Braunstein G, Shifren J, Simon J, Lucas J, Rodenberg C, Watts N. Testosterone patches for the treatment of low sexual desire in surgically menopausal women. Presentation at the 14<sup>th</sup> Annual Meeting of The North American Menopause Society; September 17-20, 2003; Miami Beach, FL.