Journal of Sex & Marital Therapy

ISSN: 0092-623X (Print) 1521-0715 (Online) Journal homepage: http://www.tandfonline.com/loi/usmt20

The Female Sexual Function Index (FSFI): A Multidimensional Self-Report Instrument for the Assessment of Female Sexual Function

R. Rosen, C. Brown, J. Heiman, S. Leiblum, C. Meston, R. Shabsigh, D. Ferguson, R. D'Agostino

To cite this article: R. Rosen, C. Brown, J. Heiman, S. Leiblum, C. Meston, R. Shabsigh, D. Ferguson, R. D'Agostino (2000) The Female Sexual Function Index (FSFI): A Multidimensional Self-Report Instrument for the Assessment of Female Sexual Function, Journal of Sex & Marital Therapy, 26:2, 191-208, DOI: 10.1080/009262300278597

To link to this article: http://dx.doi.org/10.1080/009262300278597

Journal of Sex & Marital Therapy, 26:191–208, 2000

0092-623X/00 $12.00 + .00

The Female Sexual Function Index (FSFI): A Multidimensional Self-Report Instrument for the Assessment of Female Sexual Function

R. ROSEN
UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey, USA

C. BROWN
University of Tennessee, Memphis, Tennessee, USA

J. HEIMAN
University of Washington, Seattle, Washington, USA

S. LEIBLUM
UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey, USA

C. MESTON
University of Texas, Austin, Texas, USA

R. SHABSIGH
Columbia University School of Medicine, New York, New York, USA

D. FERGUSON
Grayslake, Illinois, USA

R. D'AGOSTINO, JR.
Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA

This article presents the development of a brief, self-report measure of female sexual function. Initial face validity testing of questionnaire items, identified by an expert panel, was followed by a study aimed at further refining the questionnaire. It was administered to

Supported by Zonagen Inc. (The Woodlands, TX, USA) and Bayer AG (Wuppertal, Germany). The authors would like to thank Anita K. Islam, Ph.D. (Target Health Inc. New York, NY) for her work in the preparation of this manuscript.

Thanks also to: I. Goldstein, M.D. and J. Berman, M.D., Boston University School of Medicine, MA; J. Fourcroy, M.D., Ph.D, Rockville, MD; E. Laan, Ph.D., University of Amsterdam, The Netherlands; P. Lammers, M.D., Zonagen, Inc., The Woodlands, TX; J. Mitchel, Ph.D., J. Hays, M.S., Target Health Inc., New York, NY; H. Padma-Nathan, M.D., The Male Clinic, Santa Monica, CA; and N. Phillips, M.D., UMDNJ-Robert Wood Johnson Medical School, New Brunswick, NJ.

Address correspondence to Raymond Rosen, Ph.D., Dept. of Psychiatry, Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA.

131 normal controls and 128 age-matched subjects with female sexual arousal disorder (FSAD) at five research centers. Based on clinical interpretations of a principal components analysis, a 6-domain structure was identified, which included desire, subjective arousal, lubrication, orgasm, satisfaction, and pain. Overall test-retest reliability coefficients were high for each of the individual domains ($r = 0.79$ to 0.86) and a high degree of internal consistency was observed (Cronbach's alpha values of 0.82 and higher) Good construct validity was demonstrated by highly significant mean difference scores between the FSAD and control groups for each of the domains ($p \leq 0.001$). Additionally, divergent validity with a scale of marital satisfaction was observed. These results support the reliability and psychometric (as well as clinical) validity of the Female Sexual Function Index (FSFI) in the assessment of key dimensions of female sexual function in clinical and nonclinical samples. Our findings also suggest important gender differences in the patterning of female sexual function in comparison with similar questionnaire studies in males.

Female sexual dysfunction has traditionally included disorders of desire/libido, arousal, pain/discomfort, and inhibited orgasm. While epidemiologic data are limited, the available estimates are that 43% of women complain of at least one sexual problem (Laumann, Paik, & Rosen, 1999), while 11–33% of survey and clinical samples fall within a specific problem category (Laumann, Paik, & Rosen, 1999; Laumann, Gagnon, Michael, & Michaels, 1994; Rosen, Taylor, Leiblum, & Bachmann, 1993). Desire and arousal phase disorders are among the most common presenting problems in clinical settings. In community studies, orgasm and arousal disorders are equally prevalent (Spector & Carey, 1990). In spite of the high prevalence which appears to surpass that of male sexual dysfunction (Laumann, Paik, & Rosen, 1999), less attention has been paid to the sexual problems of women. Few studies have investigated the physiological and psychological underpinnings of female sexual dysfunction and fewer treatments are available than for comparable conditions in the male. A major barrier to the development of clinical research in this area has been the absence of well-defined endpoints and outcomes, which in turn reflects the current lack of consensus in regards to the definition and diagnostic framework for assessing and treating female sexual dysfunction. Sexual response and dysfunction are conceptualized as involving both psychological and organic processes; however, the American Psychiatric Association's Diagnostic & Statistical Manual, version IV (DSM-IV) (American Psychiatric Association, 1994) system of classification was intended as a nomenclature for mental disorders only.

Recently, an international, multi-disciplinary consensus development conference was held in the United States to develop a new classification

system to apply to all forms of female sexual dysfunction regardless of etiology (International Consensus Development Conference on Female Sexual Dysfunctions: Definitions and Classifications, in press). This panel recommended maintaining four major categories of dysfunction (desire disorders, arousal disorder, orgasmic disorder, and sexual pain disorders), as described in the DSM-IV and ICD-10 (International Classification of Diseases) (World Health Organization, 1992). At the same time, definitions of individual disorders were in some cases altered to reflect current clinical and research practices. For example, recognizing the frequent involvement of overlapping subjective and physiological aspects of female sexual arousal, the definition of Sexual Arousal Disorder was expanded to incorporate these dimensions, referring to "the persistent or recurrent inability to attain or to maintain sufficient sexual excitement, which causes personal distress. It may be expressed as lack of subjective excitement or lack of genital (lubrication/swelling) or other somatic responses." The DSM-IV and ICD-10 also recognize the need for a subjective distress criterion in defining sexual dysfunction. The assessment of personal distress is typically conducted by a clinical interview or a standardized questionnaire. However, at the present time there are no validated sexual distress specific instruments.

Laboratory-based physiological indices of sexual response (e.g., vaginal blood flow) are available. However, it has been proposed that the most valid way to assess sexual function in women is in a naturalistic setting, and the self-report technique is the only method currently available for measuring sexual response in an at-home setting. Earlier measures were primarily unidimensional scales of sexual experience and behavior (Hoon, Hoon, & Wincze, 1976; Zuckerman, 1973). More recent questionnaires provide greater detail on sexual frequency, desire, and arousal, but were developed prior to the current classification of sexual disorders and do not address some aspects of the current definitions (Derogatis & Melisaratos, 1979; Taylor, Rosen, & Leiblum, 1994).

Each of the different disorders (hypoactive sexual desire disorder, sexual arousal disorder, sexual pain disorders, and orgasmic disorder) has correspondence to different phases or domains of sexual function. Sexual satisfaction is considered one of the important dimensions of sexual function, although no separate diagnostic category is provided for dysfunction in this dimension. Sexual response involves a temporal sequencing and coordination of several phases (Masters & Johnson, 1966). Therefore, problems affecting one domain may interact with other disorders in a complex fashion, resulting in substantial overlap among diagnostic categories (Laumann, Paik, & Rosen, 1999; Rosen, Taylor, Leiblum, & Bachmann, 1993). Since differentiation of the hierarchy of disorders within a given patient may be quite difficult, instruments are needed that allow measurement of the relative strength of dysfunction within each domain. Such instruments may serve to assist and standardize diagnosis and guide treatment, as well as determining response to pharmacologic and non-pharmacologic therapies.

Unlike male sexual arousal, which is relatively easy to assess and evaluate, female sexual arousal has tended to be neglected as a target of diagnostic or empirical research. It has been difficult to specify what dimensions characterize female arousal and what constitutes an arousal disorder. Similarly, there are, at present, no empirically validated instruments for assessing female arousal, which impedes pharmacological and psychological research in female sexual response. Without an objectively validated self-report instrument that demonstrates both reliability and validity, outcome research in female sexuality will continue to lag behind that of male sexuality. The present study was undertaken in order to develop and validate a brief self-report measure of female sexual arousal and other relevant domains of sexual functioning in women.

METHODS

Phase I: Item Selection

In the initial phase of item selection, emphasis was placed on the selection of items related to Female Sexual Arousal Disorder (FSAD), although the domains of desire, pain, orgasm, and satisfaction were also addressed. A panel of experts contributed items for the questionnaire, initially consisting of 30 items that sampled each of the domains. The intent of this initial phase was to meet basic psychometric criteria, i.e., to be clear and understandable, to provide comprehensive response choices, to be relatively simple to administer and score, and to be unbiased with respect to age, ethnicity, education, or economic status. The questions were designed to be suitable for use with a heterosexual as well as a homosexual population. The questionnaire developed from this process was administered to a sample of 30 female volunteers at three investigational sites for individual critique as well as critique in a focus group setting.

The data from this initial validation study indicated that the majority of subjects understood the questions and response options, although some of the questions and choices were re-worded to improve clarity. The feedback was reviewed by the expert panel, which then made additional suggestions for modification of individual items, removal of some items, and development of new items. To be consistent with the "personal distress" criterion of current classification systems, a decision was made to address issues of sexual distress in the next phase of test validation.

Phase 2: Reliability and Construct Validity

In the second phase, the 29-item questionnaire resulting from the phase 1 testing was administered to 259 volunteer subjects recruited from the general population in a 5-center study. The centers were: UMDNJ-Robert Wood

Johnson Medical School, Piscataway, NJ; U. Tennessee, Memphis, TN; U. Washington, Seattle, WA.; U. Texas, Austin, TX; Columbia U. School of Medicine, New York, NY. The clinical sample met criteria for the clinical diagnosis of Female Sexual Arousal Disorder, while the control sample reported no problems with arousal, desire, or orgasm, and were sexually active and engaged in a stable heterosexual relationship. Females between 21 and 70 years-of-age could be enrolled and attempts were made to have participants from a wide range of ages. Subjects from the clinical sample were age-matched (± 2 years) with subjects from the nonclinical sample.

The goal of the second phase was to examine the FSFI for construct validity (factorial, discriminant, and divergent) and reliability (internal consistency and test-retest reliability). The FSFI was administered twice (Visit 1 and Visit 2), two to four weeks apart, to assess test-retest reliability. At Visit 1, patients also completed a demographic questionnaire and a measure of marital satisfaction (Locke-Wallace) to provide an assessment of divergent validity (low association with a test that measures a different domain or entity).

The first series of evaluations were performed on an item-by-item basis in order to obtain items with adequate psychometric properties and clinical relevance for the final questionnaire. Factor analysis was performed to assess the underlying domain structure of the questionnaire, and to evaluate factorial validity, i.e., the degree to which each item is related to the hypothesized domain with which it is associated. Discriminant validity, the ability of each item to differentiate between controls and FSAD subjects was assessed by using a between-groups analysis of variance. Internal consistency (relatedness) of the items within each factor was evaluated using the Cronbach's alpha formula. Test-retest reliability, the stability quotient of each item from Visit 1 to Visit 2, was determined by means of Pearson Product-moment correlations. In the final series of analyses, the psychometric properties of the questionnaire resulting from the above item reduction process were fully evaluated.

RESULTS

Study Participant Characteristics

A summary of the baseline characteristics of study participants is presented in Table 1. The mean (± SD) ages of the FSAD group (N = 128) and control group (N = l31) were 40.5 ± 12.98 years and 39.7 ± 13.15 years, respectively. Age ranges were 21 to 69 years (FSAD) and 21 to 68 years (controls). The two groups were comparable with respect to race, education, income, and number of children. There was a statistically significant difference between the two groups in marital status ($p = 0.042$), with a greater proportion of married women in the FSAD group (69% versus 55%) and more single women in the control group (34% versus 19%). The two groups did not differ signifi-

cantly in the proportion of women taking hormone replacement therapy. There was, however, a statistically significant group difference in the use of anti-depressant medications, with 19% of the subjects with FSAD currently taking some form of medication for depression, in comparison to 7% of the healthy controls (p = 0.004). The frequency of sexual activity also differentiated the two groups to a statistically significant degree, with lower frequency in the FSAD group compared to the control group. A significantly greater percentage of healthy volunteers than women with FSAD also used some form of contraceptive (54% versus 39%, respectively, p = 0.033). Detailed demographic data for the two groups are provided in Table 1.

TABLE 1. Baseline Characteristics

	FSAD	Control	P value*
	N = 128	N = 131	
Age			
Mean	40.5	39.7	0.624
Min–Max	21–69	21–68	
SD	12.98	13.15	
Race	n (%)	n (%)	0.932
Caucasian	98 (76.6)	100 (76.3)	
African-American	14 (10.9)	15 (11.5)	
Native-American	1 (0.8)	0 (0)	
Hispanic	9 (7.0)	11 (8.4)	
Asian	5 (3.9)	4 (3.1)	
Other	1 (0.8)	1 (0.8)	
Education			0.921
High school	27 (21.1)	22 (16.8)	
GED	3 (2.3)	3 (2.3)	
2 years college	41 (32.0)	43 (32.8)	
4 years college	30 (23.4)	35 (26.7)	
Graduate school	27 (21.1)	28 (21.4)	
Annual Income			0.761
< 50,000	60 (46.9)	67 (51.2)	
50,000–100,000	43 (33.6)	39 (29.8)	
>100,000	25 (19.5)	25 (19.1)	
Marital Status			0.042
Married	88 (68.8)	71 (51.6)	
Divorced	16 (12.5)	15 (11.5)	
Single	24 (18.8)	44 (33.9)	
Children (Yes)	72 (57.6)	78 (59.5)	0.753
Hormone Replacement Therapy (Yes)	33(25.8)	23 (17.6)	0.108
Anti-Depressant Use (Yes)	24 (18.8)	9 (6.9)	0.004
Frequency of Sexual Activity			0.009
< once per month	21 (16.4)	9 (6.9)	
1–2 per month	31 (24.2)	24 (18.3)	
1–2 per week	56 (43.8)	56 (42.8)	
3–4 per week	14 (10.9)	31 (23.7)	
> 4 per week	6 (4.7)	11 (8.4)	

*P values assessed using between-groups analysis of variance or Chi square/Fishers' Exact test

Factor Analysis and Domain Scoring

A principal components analysis (using varimax rotation) was performed to investigate the factor structure of the 29-item questionnaire. This analysis was performed using Visit I data for the total group (N = 259) as well as separately for the FSAD subgroup (N = 128) and controls (N = 131). Four factors with eigenvalues over 1.00 were identified for all three samples; however, a five-factor solution with the FSAD group yielded the most consistent pattern of factor loadings. The fifth factor had an eigenvalue high enough to justify inclusion.

The selection of items for each factor was determined on the basis of a combination of the following statistical and clinical considerations. Statistical criteria for item inclusion included high/moderate loading on one factor, low cross-factor loading, high/moderate test-retest reliability, and good discrimination between the FSAD and the control sample. While all of the items were satisfactory to very satisfactory on these criteria, a refinement of the full scale through item reduction was conducted for the goal of minimizing redundancy and creating a briefer instrument. Four questions relating to "pleasurable feelings" of warmth and tingling were eliminated because there could be some ambiguity in defining these terms (they also showed more factor cross-loading than other items). A question about "satisfaction with amount of stimulation from partner" was dropped since three good satisfaction items were already available. Additionally, the five compound (two-part) questions which related to "distress" were dropped from further consideration due to their dependence on sexual activity and resultant complications in interpretation and scoring. All subsequent analyses were performed on the 19-item questionnaire resulting from this item reduction process.

In order to assess the impact of choosing the response choice of "no sexual activity," a factor analysis was performed excluding data from the 13 subjects in the FSAD group and 3 subjects in the control group who reported that they had no sexual activity within the past month. Since a pattern of factor results was obtained that was very similar to that described below for the complete sample, further analyses were performed with all study subjects included.

Based on the results of the factor analysis of the 19-item questionnaire using the complete sample, individual items were assigned to five separate domains of female sexual function: desire/arousal, lubrication, orgasm, satisfaction, and pain. The pattern of factor loadings was similar for the total group and the FSAD group alone. The pattern was less clear for the control sample alone. In this group, for example, one factor was comprised of a mix of lubrication and arousal items, and another consisted of a mix of arousal and orgasm items. Table 2 presents the confirmatory factor analysis results for the FSAD group. Items generally clustered in the predicted fashion and had relatively high factor loadings, supporting the factoral validity of the final questionnaire.

TABLE 2. Principal components analysis using varimax rotation of the 19 final questions of the FSFI

Item	Factors 1	2	3	4	5
1. Desire: frequency	83*	11	15	13	7
2. Desire: level	84*	17	9	19	5
3. Arousal: frequency	63*	26	25	50*	–1
4. Arousal: level	67*	27	30	46*	–2
5. Arousal: confidence	67*	14	33	34	6
6. Arousal: satisfaction	44*	26	52*	46*	1
7. Lubrication: frequency	32	78*	17	19	11
8. Lubrication: difficulty	24	74*	29	6	26
9. Lubrication: frequency of maintaining	16	84*	11	18	22
10. Lubrication: difficulty in maintaining	3	82*	22	5	34
11. Orgasm: frequency	20	20	83*	15	–6
12. Orgasm: difficulty	20	19	84*	12	9
13. Orgasm: satisfaction	19	22	76*	41*	–1
14. Satisfaction: with amount of closeness with partner	18	8	8	62*	25
15. Satisfaction with sexual relationship	27	17	21	75*	14
16. Satisfaction: with overall sex life	33	5	30	70*	15
17. Pain: frequency *during* vaginal penetration	10	26	3	11	83*
18. Pain: frequency *following* vaginal penetration	–1	19	–5	15	85*
19. Pain: level during or following vaginal penetration	4	19	4	13	89*
Eigenvalue	8.45	2.78	1.52	1.15	0.65

*Items with the highest loadings within each factor.

While the statistical solution supported a 5-factor solution, clinical consideration dictated separating the domain of desire/arousal. Therefore, the final set of items was assigned to a six domain instrument measuring: (a) desire, (b) arousal, (c) lubrication, (d) orgasm, (e) global satisfaction, and (f) pain. Questions relating to domain-specific aspects of satisfaction, which tended to have a moderate loading on a specific domain as well on the overall satisfaction factor, were assigned to the specific domain. Therefore, the satisfaction domain is comprised only of items pertaining to global sexual and relationship satisfaction. This can also be viewed as the "quality of life" domain of the scale.

A simple computational algorithm was devised for determining domain scores and a composite full-scale score (see Appendices A and B). The range of possible domain scores is presented in Table 3.

As indicated in Table 4, the domain intercorrelations were generally lower for the FSAD group than for the controls or total group. The correlations between pain and desire (r = 0. 15) and between pain and orgasm (r = 0.10) failed to reach statistical significance for the FSAD group. For FSAD subjects, the highest positive correlation was between the domains of desire and arousal (r = 0.71), consistent with the factor analysis results described above.

TABLE 3. Domain Scoring

Domain	Item Number	Score Range	Minimum Score	Maximum Score
Desire	1, 2	1–5	2	10
Arousal	3, 4, 5, 6	0–5	0	20
Lubrication	7, 8, 9, 10	0–5	0	20
Orgasm	11, 12, 13	0–5	0	15
Satisfaction	14, 15, 16	0 (or 1)–5*	2	15
Pain	17, 18, 19	0–5	0	15

*Range for item 14 = 0–5; range for items 15 and 16 = 1–5

Reliability

Two kinds of test reliability were assessed: internal consistency (relatedness of items within a factor) and test-retest reliability (stability of responses over time). Using the Cronbach's alpha statistic, internal consistency was determined separately for the six domains as well as for all of the individual items (Table 5). Test results from the first administration were used and the analyses were conducted for the full sample, as well as for each subgroup (control and FSAD groups) separately. Reliability was determined for each of the domains and for the full scale score. As shown, high inter-item correlations

TABLE 4. Domain Intercorrelations (Pearson r: range = –1.00 – +1.00)

Total Group	D	A	L	O	S	P
D	1.00					
A	0.76	1.00				
L	0.56	0.75	1.00			
O	0.54	0.81	0.68	1.00		
S	0.60	0.80	0.62	0.70	1.00	
P	0.37	0.47	0.64	0.41	0.53	1.00
FSAD Group	**D**	**A**	**L**	**O**	**S**	**P**
D	1.00					
A	0.71	1.00				
L	0.38	0.51	1.00			
O	0.37	0.65	0.46	1.00		
S	0.46	0.68	0.37	0.46	1.00	
P	0.15*	0.18	0.47	0.10*	0.34	1.00
Control Group	**D**	**A**	**L**	**O**	**S**	**P**
D	1.00					
A	0.61	1.00				
L	0.40	0.70	1.00			
O	0.34	0.74	0.52	1.00		
S	0.46	0.70	0.50	0.66	1.00	
P	0.32	0.49	0.66	0.40	0.44	1.00

KEY: D = Desire; A = Arousal; L = Lubrication; O = Orgasm; S = Satisfaction; P = Pain
* Not statistically significant. All other correlation coefficients were statistically significant, $p \leq 0.05$

TABLE 5. FSFI Domain Characteristics: Reliability

	Internal Consistency*		
	Full Sample	FSAD patients	Controls
Domain			
Desire	0.92	0.91	0.89
Arousal	0.95	0.92	0.90
Lubrication	0.96	0.93	0.95
Orgasm	0.94	0.92	0.91
Satisfaction	0.89	0.82	0.91
Pain	0.94	0.93	0.92
All Items	0.97	0.93	0.95

	Test-Retest Reliability					
	Full Sample		FSAD patients		Controls	
Domain	N	r**	N	r**	N	r**
Desire	202	0.83	99	0.80	103	0.77
Arousal	200	0.85	97	0.68	103	0.85
Lubrication	201	0.86	98	0.71	103	0.89
Orgasm	200	0.80	99	0.62	101	0.87
Satisfaction	200	0.83	98	0.70	102	0.82
Pain	203	0.79	100	0.69	103	0.87
Total Scale	198	0.88	97	0.70	101	0.91

*Cronbach's alpha (range = –1.00 – +1.00).
**Pearson product-moment correlation coefficient (range = –1.00, – +1.00). All are significant, $p \leq 0.001$

were observed for all six domains (Cronbach's alpha values of 0.82 and higher).

Test-retest reliability was assessed by computing the stability coefficient (Pearson product-moment correlation) between Visits 1 and 2 scores. As seen in Table 5, overall test-retest reliability was relatively high for all of the domains (r = 0.79 – 0.86) and for the total scale (r = 0.88). For the FSAD group, the domain of desire showed the highest test-retest reliability (r = 0.80), with the other domains showing moderately high correlations (r = 0.62 – 0.71). In general, higher reliability of domain scores was obtained for the control group than for the FSAD group, though all reliability coefficients were statistically significant.

Discriminant Validity

Discriminant validity (the ability of the scale to differentiate between clinical and nonclinical populations), was assessed by comparing the mean responses of patients with FSAD with those of the controls. Table 6 presents the means of each domain and the individual items within the domains. As shown, statistically significant differences between the groups were observed for all the domains and for the full scale score. The largest differences between the groups were seen for the domains of lubrication and arousal.

TABLE 6. FSFI Domain Characteristics: Discriminant Validity

	FSAD patients		Controls		
	N	Mean ± SD	N	mean ± SD	P value*
Domains and Domain Items					
Desire	127	4.7 ± 2.12	131	6.9 ± 1.89	≤0.001
1. Desire: frequency		2.4 ± 1.13		3.4 ± 1.04	
2. Desire: level		2.3 ± 1.09		3.5 ± 0.96	
Arousal	127	9.7 ± 4.78	130	16.8 ± 3.62	≤0.001
3. Arousal: frequency		2.6 ± 1.47		4.4 ± 1.06	
4. Arousal: level		2.5 ± 1.27		4.0 ± 1.01	
5. Arousal: confidence		2.5 ± 1.25		4.1 ± 1.06	
6. Arousal: satisfaction		2.1 ± 1.30		4.4 ± 1.01	
Lubrication	127	10.9 ± 5.48	130	18.6 ± 3.17	≤0.001
7. Lubrication: frequency		2.6 ± 1.48		4.6 ± 0.91	
8. Lubrication: difficulty		2.8 ± 1.46		4.7 ± 0.79	
9. Lubrication: frequency of maintaining		2.5 ± 1.54		4.6 ± 0.92	
10. Lubrication: difficulty in		3.0 ± 1.56		4.7 ± 0.79	
Orgasm	128	7.1 ± 4.08	129	12.7 ± 3.16	≤0.001
11. Orgasm: frequency		2.4 ± 1.54		4.1 ± 1.21	
12. Orgasm: difficulty		2.5 ± 1.47		4.3 ± 1.11	
13. Orgasm: satisfaction		2.2 ± 1.40		4.4 ± 1.11	
Satisfaction	127	8.2 ± 3.59	130	12.8 ± 3.03	≤0.001
14. Satisfaction: with amount of closeness with partner		3.4 ± 1.57		4.3 ± 1.12	
15. Satisfaction: with a sexual relationship		2.6 ± 1.37		4.2 ± 1.06	
16. Satisfaction: with overall sex life		2.3 ± 1.22		4.2 ± 1.11	
Pain	128	10.1 ± 4.64	130	13.9 ± 2.79	≤0.001
17. Pain: frequency *during* vaginal penetration		3.2 ± 1.70		4.5 ± 1.09	
18. Pain: frequency following vaginal penetration		3.5 ± 1.72		4.7 ± 0.98	
19. Pain: level during or following vaginal penetration		3.4 ± 1.54		4.7 ± 0.91	
Full Scale	126	19.2 ± 6.63	129	30.5 ± 5.29	≤0.001

*P values for domain scores assessed using between-groups analysis of variance

Divergent Validty

Another method of establishing that the instrument is specifically measuring the construct under study (i.e., sexual function) is to show statistical dissociation (incomplete overlap) from an instrument that assesses a different, albeit partially related, construct (e.g., marital satisfaction). The degree of association between the FSFI scores (domains and full scale) and the Locke-Wallace Marital Adjustment Test score was calculated by means of the Pearson product-moment correlation (Table 7 below). These correlations were performed on a sample of 219 subjects (103 FSAD; 116 controls) for whom data were available on every item of the Locke-Wallace.

TABLE 7. FSFI domain characteristics: divergent validity

	Full Sample		FSAD Patients		Controls	
Domain	Pearson r	P value	Pearson r	P value	Pearson r	P value
Desire	0.19	0.005	0.04	0.714	0.16	0.086
Arousal	0.37	≤0.001	0.19	0.059	0.43	≤0.001
Lubrication	0.30	≤0.001	0.09	0.374	0.42	≤0.001
Orgasm	0.27	≤0.001	0.03	0.767	0.37	≤0.001
Satisfaction	0.57	≤0.001	0.40	≤0.001	0.72	≤0.001
Pain	0.32	≤0.001	0.20	0.039	0.33	≤0.001
Full Scale	0.41	≤0.001	0.22	0.027	0.53	<0.001

KEY: D = Desire; A = Arousal; L = Lubrication; O = Orgasm; S = Satisfaction; P = Pain

As further support for the construct validity of the FSFI, the correlations between the Locke-Wallace and the FSFI, even when statistically significant, were generally modest in magnitude. The correlation between the Locke-Wallace and the total FSFI score was low moderate for the control group (r = 0.53) and very low for the FSAD group (r = 0.22). In both groups, the strongest overlap with the marital adjustment test was observed for the satisfaction domain of the FSFI. In the FSAD group, the domains of orgasm and sexual desire showed the least association with marital adjustment. The domain of desire also showed low correlation with marital adjustment in the control group.

DISCUSSION

The objective of the present study was to develop a brief, valid, and reliable self-report measure of female sexual function, which could be easily administered to women across a wide age range, including post-menopausal women. The FSFI was designed to be a clinical trials assessment instrument that addresses the multidimensional nature of female sexual function.

The FSFI was developed in a series of stages, including panel selection of the initial items, pre-testing with healthy volunteers followed by linguistic and conceptual validation with a panel of expert consultants. Based on factor analytic methods, five factors or domains of sexual function were identified: (a) desire and subjective arousal, (b) lubrication, (c) orgasm, (d) satisfaction, and (e) pain/discomfort. The factor loadings of the individual items fit the expected pattern, supporting the factorial validity of this instrument.

When the first factor of desire and arousal is considered as two separate domains, the strength of the relationship between these two domains is greater than that between the other domains (r = 0.76). This relationship demonstrates a considerable overlap between the dimensions of desire and arousal in women, consistent with clinical observation and contrasting with findings in studies of sexual dysfunction in men (Rosen, Riley, Wagner, Osterloh, Kirkpatrick, & Mishra, 1997). Nevertheless, desire and arousal can

be defined independently, and in consideration of recent recommendations for a new classification system of female sexual dysfunction, a clinically-based decision was made to separate the mixed factor of desire/arousal into two measurable dimensions. The advantage of this new scale is that it includes measures of both peripheral (e.g., lubrication) and central (subjective arousal and desire, as separable dimensions) responses to sexual stimulation. A panel of experts concluded that the inclusion of such clinical endpoints would provide greater ability to assess treatment specificity.

After a process of item reduction, 19 items with optimal psychometric properties, sampled from each of six domains, were selected. A scoring system was developed for obtaining individual domain scores and a full scale score. Psychometric validation of the 19-item questionnaire was conducted in two major areas: (a) construct validity and (b) test reliability. Specifically, individual domain scores of the FSFI, as well as the full scale score, showed good ability to differentiate between FSAD and non-FSAD subjects (discriminant validity), while showing only modest correlations with a measure of marital satisfaction (divergent validity). Divergence of the FSFI from the Locke-Wallace Marital Adjustment Test was greater for the FSAD group than for the controls, for whom a moderately high correlation was observed between marital satisfaction and the FSFI domain of global sexual satisfaction. For subjects with sexual dysfunction, FSFI scores appeared to be fairly independent of the influence of marital adjustment. The individual domains of the FSFI and the FSFI full scale score showed high internal consistency and acceptable test-retest reliability. The domain intercorrelations were generally lower for the FSAD group than for the controls, suggesting a greater level of independence among the dimensions of sexual function in patients with FSAD than in women with normal sexual response.

There were some differences between the FSAD and control groups in demographic factors. A significantly greater number of FSAD than control subjects (19% vs 7%) were taking antidepressant medications, which may have contributed to drug-induced sexual dysfunction. However, the intent of the study was to discriminate between sexual dysfunction and control populations, rather than to differentiate between specific etiologies of sexual dysfunction. In addition, significantly more controls than FSAD subjects (34% vs 19%) were unmarried. A recent study showed that unmarried women are 1½ times more likely to have climax problems (Laumann, Paik, & Rosen, 1999) a factor that may have reduced the ability to discriminate between groups on the orgasm domain.

Although a response to some of the questions was dependent on whether the subject engaged in sexual activity, in the current study, subjects could have been selected who were not sexually active. Nevertheless, interpretation of the study results were not affected, since a reanalysis of the FSFI factor structure following the exclusion of data from subjects who reported that they had no recent sexual activity showed little change. However, to derive an unambiguous full scale score, the FSFI is appropriately used only

for subjects who have had some level of sexual activity during the measurement period. Also, a score of zero in a domain total indicates that no sexual activity has been reported, but conveys no further meaningful information about the specific domain.

Dependence on sexual activity did affect the original items relating to psychological distress, which were not included in the final questionnaire due to ambiguity in their interpretation. However, it should be noted, that while distress is an important component in the diagnosis of female sexual dysfunction, the evaluation of pharmaceutical agents will focus on perceptions of sexual responsiveness (e.g., level of arousal and lubrication) and global satisfaction as the clinical outcomes. The specific issue of distress in the diagnosis of female sexual dysfunction is currently being examined by a consensus panel of the American Foundation of Urological Diseases, and will be the subject of a future publication.

CONCLUSIONS

The FSFI, a 19-item questionnaire, has been developed as a brief, multidimensional self-report instrument for assessing the key dimensions of sexual function in women. It is psychometrically sound, easy to administer, and has demonstrated ability to discriminate between clinical and nonclinical populations. The questionnaire described was designed and validated for assessment of female sexual function and quality of life in clinical trials or epidemiological studies. Its further use in these areas remains to be investigated.

REFERENCES

American Psychiatric Association. (1994). *DSM-IV: Diagnostic and statistical manual of mental disorders* (4th ed., pp. 493–518). Washington, DC: American Psychiatric Association.

Derogatis, L. R., & Melisaratos, N. (1979). The DSFI: a multidimensional measure of sexual functioning. *J. Sex Marital Ther. 5*: 244–281.

Hoon, E. F., Hoon, P. W., & Wincze, J. P. (1976). An inventory for the measurement of female sexual arousability. *Arch. Sex. Behav. 5*: 291–300.

International Consensus Development Conference on Female Sexual Dysfunctions: Definitions and Classifications (in press). Boston, MA 1998. *J. Urology*.

Laumann, E. O., Paik, A., & Rosen, R. C. (1999). Sexual dysfunction in the United States. *JAMA 281*: 537–544.

Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The Social Organization of Sexuality*. Chicago: University of Chicago Press.

Masters, W. H., & Johnson, V. E. (1966). *Human sexual response*. Boston: Little Brown.

Rosen, R. C., Riley, A., Wagner, G., Osterloh, I. H., Kirkpatrick, J., & Mishra, A. (1997). The International Index of Erectile Function (IIEF): a multidimensional scale for assessment of erectile dysfunction. *Urology 49*: 822–830.

Rosen, R. C., Taylor, J. F., Leiblum, S. R., & Bachmann, G. A. (1993). Prevalence of sexual dysfunction in women: results of a survey of 329 women in an outpatient gynecological clinic. *Journal of Sex & Marital Therapy, 19*: 171–188.

Spector, I. P., & Carey, M. P. (1990). Incidence and prevalence of the sexual dysfunctions: a critical review of the empirical literature. *Arch. Sex. Behav, 19*: 389–408.

Taylor, J. F., Rosen, R. C., & Leiblum, S. R. (1994). Self-report assessment of female sexual function: psychometric evaluation of the brief index of sexual functioning for women. *Arch. Sex. Behav. 23*: 627–643.

World Health Organization. (1992). ICD-10: International Statistical Classification of Diseases and Related Health Problems. World Health Organization: Geneva.

Zuckerman, M. (1973). Scales for sex experience for males and females. *J. Consul. Clin Psychol. 41*: 27–29.

Appendix A—Female Sexual Function Index (FSFI)*

Question	Response Options
Q1: Over the past 4 weeks, how **often** did you feel sexual desire or interest?	5 = Almost always or always 4 = Most times (more than half the time) 3 = Sometimes (about half the time) 2 = A few times (less than half the time) 1 = Almost never or never
Q2: Over the past 4 weeks, how would you rate your **level** (degree) of sexual desire or interest?	5 = Very high 4 = High 3 = Moderate 2 = Low 1 = Very low or none at all
Q3. Over the past 4 weeks, how **often** did you feel sexually aroused ("turned on") during sexual activity or intercourse?	0 = No sexual activity 5 = Almost always or always 4 = Most times (more than half the time) 3 = Sometimes (about half the time) 2 = A few times (less than half the time) 1 = Almost never or never
Q4. Over the past 4 weeks, how would you rate your **level** of sexual arousal ("turn on") during sexual activity or intercourse?	0 = No sexual activity 5 = Very high 4 = High 3 = Moderate 2 = Low 1 = Very low or none at all
Q5. Over the past 4 weeks, how **confident** were you about becoming sexually aroused during sexual activity or intercourse?	0 = No sexual activity 5 = Very high confidence 4 = High confidence 3 = Moderate confidence 2 = Low confidence 1 = Very low or no confidence
Q6. Over the past 4 weeks, how **often** have you been satisfied with your arousal (excitement) during sexual activity or intercourse? Response Options	0 = No sexual activity 5 = Almost always or always 4 = Most times (more than half the time) 3 = Sometimes (about half the time) 2 = A few times (less than half the time) 1 = Almost never or never

Question	Response Options
Q7: Over the past 4 weeks, how **often** did you become lubricated ("wet") during sexual activity or intercourse?	0 = No sexual activity 5 = Almost always or always 4 = Most times (more than half the time) 3 = Sometimes (about half the time) 2 = A few times (less than half the time) 1 = Almost never or never
Q8. Over the past 4 weeks, how **difficult** was it to become lubricated ("wet") during sexual activity or intercourse?	0 = No sexual activity 1 = Extremely difficult or impossible 2 = Very difficult 3 = Difficult 4 = Slightly difficult 5 = Not difficult
Q9: Over the past 4 weeks, how often did you **maintain** your lubrication ("wetness") until completion of sexual activity or intercourse?	0 = No sexual activity 5 = Almost always or always 4 = Most times (more than half the time) 3 = Sometimes (about half the time) 2 = A few times (less than half the time) 1 = Almost never or never
Q10: Over the past 4 weeks, how **difficult** was it to maintain your lubrication ("wetness") until completion of sexual activity or intercourse?	0 = No sexual activity 1 = Extremely difficult or impossible 2 = Very difficult 3 = Difficult 4 = Slightly difficult 5 = Not difficult
Q11. Over the past 4 weeks, when you had sexual stimulation or intercourse, how **often** did you reach orgasm (climax)?	0 = No sexual activity 5 = Almost always or always 4 = Most times (more than half the time) 3 = Sometimes (about half the time) 2 = A few times (less than half the time) 1 = Almost never or never
Q12: Over the past 4 weeks, when you had sexual stimulation or intercourse, how **difficult** was it for you to reach orgasm (climax)?	0 = No sexual activity 1 = Extremely difficult or impossible 2 = Very difficult 3 = Difficult 4 = Slightly difficult 5 = Not difficult
Q13: Over the past 4 weeks, how **satisfied** were you with your ability to reach orgasm (climax) during sexual activity or intercourse?	0 = No sexual activity 5 = Very satisfied 4 4 = Moderately satisfied 3 = About equally satisfied and dissatisfied 2 = Moderately dissatisfied 1 = Very dissatisfied
Q14: Over the past 4 weeks, how **satisfied** have you been with the amount of emotional closeness during sexual activity between you and your partner?	0 = No sexual activity 5 = Very satisfied 4 = Moderately satisfied 3 = About equally satisfied and dissatisfied 2 = Moderately dissatisfied 1 = Very dissatisfied

Question	Response Options
Q15: Over the past 4 weeks, how satisfied have you been with your sexual relationship with your partner?	5 = Very satisfied 4 = Moderately satisfied 3 = About equally satisfied and dissatisfied 2 = Moderately dissatisfied 1 = Very dissatisfied
Q16: Over the past 4 weeks, how **satisfied** have you been with your overall sexual life?	5 = Very satisfied 4 = Moderately satisfied 3 = About equally satisfied and dissatisfied 2 = Moderately dissatisfied 1 = Very dissatisfied
Q17: Over the past 4 weeks, how **often** did you experience discomfort or pain during vaginal penetration?	0 = Did not attempt intercourse I = Almost always or always 2 = Most times (more than half the time) 3 = Sometimes (about half the time) 4 = A few times (less than half the time) 5 = Almost never or never
Q18: Over the past 4 weeks, how **often** did you experience discomfort or pain following vaginal penetration?	0 = Did not attempt intercourse 1 = Almost always or always 2 = Most times (more than half the time) 3 = Sometimes (about half the time) 4 = A few times (less than half the time) 5 = Almost never or never
Q19. Over the past 4 weeks, how would you rate your **level** (degree) of discomfort or pain during or following vaginal penetration?	0 = Did not attempt intercourse 1 = Very high 2 = High 3 = Moderate 4 = Low 5 = Very low or none at all

* For the complete FSFI questionnaire, instructions and scoring algorithm, please see www.FSFIquestionnaire.com, or contact Raymond Rosen Ph.D., (Department of Psychiatry: UMDNJ-Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854)

Appendix B—Scoring System

The individual domain scores and full scale score of the FSFI are derived by the computational formula outlined in the table below. Individual domain scores are obtained by adding the scores of the individual items that comprise the domain and multiplying the sum by the domain factor (see below). The full scale score is obtained by adding the six domain scores. It should be noted that within the individual domains, a domain score of zero indicates that no sexual activity was reported during the past month.

Domain	Questions	Score Range	Factor	Minimum score	Maximum score
Desire	1, 2	1–5	0.6	1.2	6.0
Arousal	3, 4, 5, 6	0–5	0.3	0	6.0
Lubrication	7, 8, 9, 10	0–5	0.3	0	6.0
Orgasm	11, 12, 13	0–5	0.4	0	6.0
Satisfaction	14, 15, 16	0 (or 1)–5	0.4	0	6.0
Pain	17, 18, 19	0–5	0.4	0	6.0
		Full Scale Score Range		2.0	36.0