RAPID COMMUNICATION

ELSEVIER

THE INTERNATIONAL INDEX OF ERECTILE FUNCTION (IIEF): A MULTIDIMENSIONAL SCALE FOR ASSESSMENT OF ERECTILE DYSFUNCTION

RAYMOND C. ROSEN, ALAN RILEY, GORM WAGNER, IAN H. OSTERLOH, JOHN KIRKPATRICK, AND AVANISH MISHRA

ABSTRACT

Objectives. To develop a brief, reliable, self-administered measure of erectile function that is cross-culturally valid and psychometrically sound, with the sensitivity and specificity for detecting treatment-related changes in patients with erectile dysfunction.

Methods. Relevant domains of sexual function across various cultures were identified via a literature search of existing questionnaires and interviews of male patients with erectile dysfunction and of their partners. An initial questionnaire was administered to patients with erectile dysfunction, with results reviewed by an international panel of experts. Following linguistic validation in 10 languages, the final 15-item questionnaire, the International Index of Erectile Function (IIEF), was examined for sensitivity, specificity, reliability (internal consistency and test-retest repeatability), and construct (concurrent, convergent, and discriminant) validity.

Results. A principal components analysis identified five factors (that is, erectile function, orgasmic function, sexual desire, intercourse satisfaction, and overall satisfaction) with eigenvalues greater than 1.0. A high degree of internal consistency was observed for each of the five domains and for the total scale (Cronbach's alpha values of 0.73 and higher and 0.91 and higher, respectively) in the populations studied. Test-retest repeatability correlation coefficients for the five domain scores were highly significant. The IIEF demonstrated adequate construct validity, and all five domains showed a high degree of sensitivity and specificity to the effects of treatment. Significant (P values = 0.0001) changes between baseline and post-treatment scores were observed across all five domains in the treatment responder cohort, but not in the treatment nonresponder cohort.

Conclusions. The IIEF addresses the relevant domains of male sexual function (that is, erectile function, orgasmic function, sexual desire, intercourse satisfaction, and overall satisfaction), is psychometrically sound, and has been linguistically validated in 10 languages. This questionnaire is readily self-administered in research or clinical settings. The IIEF demonstrates the sensitivity and specificity for detecting treatment-related changes in patients with erectile dysfunction.

UROLOGY **49:** 822–830, 1997.

Erectile dysfunction (ED), defined by a National Institutes of Health (NIH) Consensus Development Conference as the inability to achieve or maintain an erection sufficient for satisfactory sexual performance,[1] is estimated to affect as many as 30 million men in the United States.[2] The problem is strongly age-related, with an approximately two-fold to threefold increase in the prevalence of moderate-to-severe ED between the ages of 40 and 70 years.[2] A variety of medical, psychologic, and lifestyle factors have been implicated in the etiology of ED,[2–4] which impacts negatively on self-esteem, quality of life, and interpersonal relationships.[1]

Although laboratory-based diagnostic procedures are available, it has been proposed that sexual function is best assessed in a naturalistic setting with patient self-report techniques.[5,6] For this

This research was supported by a grant from Pfizer Inc.

From the Center for Sex and Marital Health, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey; Human Sexuality Unit, Harewood House, Springfield University Hospital, London, United Kingdom; Department of Medical Physiology, Panum Institute, University of Copenhagen, Copenhagen, Denmark; and Pfizer Central Research, Sandwich, United Kingdom, and Pfizer Central Research, Groton, Connecticut.

Reprint requests: Raymond C. Rosen, Ph.D., Department of Psychiatry, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854

Submitted: December 11, 1996, accepted (with revisions):

sensitive than unidimensional scales in the evaluation of treatment outcomes, and they are more psychometrically valid.[7] Multidimensional scales also provide greater potential for use in a clinical setting. Self-report methods are preferable to patient interview techniques, particularly in multicenter, multinational clinical trials.

Existing self-report measures of male sexual function[8–11] have several limitations, including excessive length or complexity, unacceptable patient burden, an overly narrow or restrictive focus, and inadequate psychometric, cultural, or linguistic validation. None of the current measures has been demonstrated to have adequate discriminant validity or to provide sufficient sensitivity in evaluating treatment outcomes in multinational clinical trials. Additionally, factor analytic methods were not used in the development of existing measures. Despite these limitations, self-report measures provide essential data on male sexual function in both research and clinical settings.[5] A strong recommendation of the NIH Consensus Conference was to develop better and more reliable methods for assessing the symptoms of ED and relevant treatment outcomes.[1]

The objective of the present research was to develop a brief and reliable measure of erectile function that is culturally, linguistically, and psychometrically valid. State-of-the-art methods for questionnaire development were used, and a multidimensional measure was designed to provide sensitive and specific outcome assessments in clinical trials of ED. Finally, the goal was to develop a self-administered questionnaire that would be suitable for use by clinicians and researchers, one that would be minimally burdensome to patients.

METHODS

Phase 1: Item Selection

Using multiple sources, relevant domains of male sexual function were identified across various cultures. A comprehensive review of the literature was conducted, and existing questionnaire instruments were evaluated. Detailed interviews of male patients with ED (n = 37) and their partners (n = 7) were also conducted in five countries. In this phase, four dimensions of male sexual function were identified: erectile function, orgasmic function, sexual desire, and sexual satisfaction. In a phase II trial of 351 patients with ED, an initial version of the questionnaire was administered and found to have a high degree of internal consistency among items (Cronbach's alpha statistic[12] greater than 0.85) and excellent treatment sensitivity ($P < 0.01$).[13] An exploratory factor analysis was performed that indicated a robust factor structure.[13,14] The results were reviewed by an international panel of experts who made recommendations for item modification and the development of additional items.

Phase 2: Cultural and Linguistic Evaluation

International Index of Erectile Function (IIEF) questionnaire in less than 15 minutes and reported little or no difficulty in comprehending the items. Linguistic validation of the instrument was conducted in 10 languages (Danish, Dutch, English [American, Australian, and British], Finnish, French, German, Italian, Norwegian, Spanish, and Swedish)* in 12 countries by the MAPI Research Institute in Lyon, France. This process included forward and back translations of the items and comprehensive testing of the final item pool. International harmonization techniques were used to ensure cross-cultural equivalence of the items in the targeted languages.

Phase 3: Reliability, Construct Validity, and Treatment Responsiveness

The final 15-item questionnaire (see Appendix) was administered in a large-scale clinical trial of patients with ED (study A), a comparison group of functional, age-matched volunteers (study B), and a clinical validation study that included both patients with ED and normal volunteers (study C). The designs of the studies and subject characteristics are summarized in Table I. Each study protocol was approved by the institutional review board at the participating site. All participants in the studies gave written informed consent. Men aged 18 years or older with a clinical diagnosis of ED of broad-spectrum etiology and of at least 6 months' duration (studies A and C) or normal volunteers (studies B and C) were eligible for enrollment. Patients with penile anatomic defects, uncontrolled major medical illnesses or psychologic disorders, or known drug or alcohol dependence were excluded from the studies.

Study A. This study consisted of a 2 to 4-week run-in phase, followed by a 12-week, double-blind, placebo-controlled phase in which 111 patients with ED of broad-spectrum etiology were randomized to receive either placebo or 25 mg (one capsule) of sildenafil (VIAGRA; Pfizer Inc.). Sildenafil is an oral medication that is being evaluated for the treatment of ED.[15,16] The placebo or sildenafil dose could be increased to 50 mg (two capsules) and then to 100 mg (four capsules) if a patient's response was suboptimal. The IIEF was self-administered at the screening visit (week −4 or −2), at the end of the run-in phase (week 0), and at the end of 2, 4, 8, and 12 weeks of double-blind treatment. A global efficacy question ("Did the treatment improve your erections?") was asked at the end of the double-blind treatment phase. The sensitivity, specificity, and reliability (internal consistency and test-retest repeatability) of the 15-item questionnaire were determined as follows. Each patient was designated as a "responder" or "nonresponder," based on his response to the end-of-treatment global efficacy question. Within each cohort, the mean and median baseline-to-end point changes in response values for each question were calculated. The sensitivity of the IIEF was assessed by evaluating the clinical relevance and statistical significance of the changes in the responder cohort. Specificity was assessed in the same manner in the nonresponder cohort. Internal consistency was evaluated by calculating Cronbach's alpha statistic on the item domains and the total scale.[12]

Study B. This study assessed the response to the IIEF questionnaire in 109 male volunteers (controls) without any history of male ED. These volunteers were age-matched to the patients randomized in study A (Table I). The IIEF was self-administered, with the results in these controls compared with those obtained in men with ED in study A using be-

* *Additional validation studies of other languages (for example,*

TABLE I. *Study designs and baseline characteristics of individuals enrolled in validation studies*

Study Design	Study A (Patients with ED)	Study B (Controls)	Study C	
			Patients with ED	Controls
Treatments	Sildenafil (25, 50, or 100 mg) or placebo	None	None	
Duration of study	12 weeks	1 day	4 weeks	
Timing of IIEF self-administration	Week −4 or −2, 0, 2, 4, 8, and 12	Day 1	Week 0 and 4	
Other relevant assessments	Global efficacy question: final visit		Clinical interview: Week 0 and 4 Locke-Wallace Scale: Week 0 Marlowe-Crowne Scale: Week 0	
Patient characteristics				
n	111	109	37	21
Mean age, yr (range)	56 (29–89)	55 (29–76)	53 (29–71)	58 (37–76)
Mean duration of ED, yr (range)	4.61 (1–37)	—	5.9 (1–18)	—
Primary etiology*				
Organic	21%	—	14%	—
Psychogenic	40%	—	49%	—
Mixed	37%	—	38%	—
Unknown	3%	—	0%	—

Key: ED = erectile dysfunction; IIEF = International Index of Erectile Function.
** Percentages do not total 100 due to rounding.*

tween-groups discriminant analysis (analysis of covariance controlling for age) and post hoc comparison of group differences on individual items.

Study C. This 4-week study evaluated the construct validity and test-retest repeatability of the IIEF in 37 patients with male ED and in 21 age-matched controls (Table I). The IIEF was self-administered at week 0 and week 4. In this study, blinded clinical interviews of patients were conducted at week 0 to evaluate the convergent validity of the measure (that is, concordance with an independent method of assessment). In addition, patients completed measures of marital satisfaction (Locke-Wallace scale[17]) and social desirability (Marlowe-Crowne scale[18]) to assess divergent validity (that is, separateness from overlapping or related constructs) at week 0. Test-retest reliability of the total and individual item scores of the IIEF were assessed by calculating the Pearson product-moment correlation coefficient[19] for each group (patients and controls). Internal consistency was evaluated using the Kuder-Richardson formula. Discriminant validity was assessed using repeated-measures analysis of variance, with subject group as the between-groups variable, time (week 0 and week 4) as the repeated-measures factor, and study measure as the outcome variable.

RESULTS

Factor Analysis and Domain Scoring

A principal components analysis (with varimax rotation) was performed to investigate the factor structure of the final 15-item questionnaire (see Appendix). Five factors with eigenvalues† greater than 1.0 were identified (Table II). Final item selection for each factor was based on a combination of statistical and clinical considerations.[20] Based on results of the confirmatory factor analysis, together with clinical interviews and expert panel consultation, the responses to individual items of the questionnaire were assigned to five separate domains of sexual function: (1) erectile function, (2) orgasmic function, (3) sexual desire, (4) intercourse satisfaction, and (5) overall satisfaction. Domain scores were computed by summing the scores for individual items in each domain. The system of domain scoring and resulting interdomain correlations are presented in Table III.

Scale Reliability

Two separate aspects of scale reliability were evaluated, namely, internal consistency and test-retest repeatability. Internal consistency (Cronbach's alpha) was computed separately for the five domains and for all items combined in each of the three test samples. Responses in the erectile and orgasmic function domains were highly consistent, with alpha values greater than 0.90 (Table IV). A satisfactory degree of consistency also was observed for items in the other domains (alpha values greater than 0.70) and for the total scale (alpha values greater than 0.90) in each of the test samples.

Test-retest repeatability was assessed in study C

TABLE II. *Principal components analysis with varimax rotation of 15 questions of International Index of Erectile Function: factor loadings**

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1. Erection frequency	**0.77**	0.03	0.31	0.17	−0.05
2. Erection firmness	**0.92**	0.12	0.20	0.08	0.04
3. Penetration ability	**0.89**	0.16	0.15	0.06	0.14
4. Maintenance frequency	**0.82**	0.26	0.13	−0.02	0.22
5. Maintenance ability	**0.68**	0.39	0.09	0.07	0.41
6. Intercourse frequency	0.10	−0.02	0.11	0.34	**0.79**
7. Intercourse satisfaction	0.61	0.28	0.31	−0.13	**0.48**
8. Intercourse enjoyment	0.53	0.39	0.18	0.01	**0.53**
9. Ejaculation frequency	0.26	0.20	**0.89**	0.10	0.13
10. Orgasm frequency	0.23	0.25	**0.87**	0.18	0.12
11. Desire frequency	0.06	−0.01	0.15	**0.88**	0.16
12. Desire level	0.04	0.26	0.07	**0.87**	0.08
13. Overall satisfaction	0.29	**0.76**	0.28	0.15	−0.01
14. Relationship satisfaction	0.18	**0.83**	0.21	0.14	0.13
15. Erection confidence	**0.65**	0.53	0.01	0.01	0.07
Eigenvalue	4.72	2.22	2.03	1.81	1.47

* *Items with the highest loadings within each factor are boldfaced.*

4 visits. As shown in Table IV, test-retest repeatability was relatively high for the erectile function ($r = 0.84$) and intercourse satisfaction ($r = 0.81$) domains, as well as for the total scale scores ($r = 0.82$). Moderately high correlations were observed for the other domains (r values of 0.64 to 0.77).

Discriminant Validity

Discriminant validity, or the ability of the IIEF scale to discriminate reliably between clinical and nonclinical populations, was assessed by comparing the responses from patients with ED with those from controls in two studies. As shown in Table V, highly significant differences were observed between the the patients with ED and age-matched controls for most domains. Differences between domain scores between these two groups were greatest for the erectile function domain ($P \leq 0.0001$), followed by intercourse satisfaction ($P \leq 0.001$) and overall satisfaction ($P \leq 0.001$). The least degree of difference between patients and controls was seen for the sexual desire domain, with results failing to reach statistical significance in study C. This result is not surprising because all patients were recruited for a clinical trial of ED and were excluded for concomitant sexual disorders, such as hypoactive sexual desire.

Convergent and Divergent Validity

To demonstrate construct validity of a new measure, it is important to show that scale scores are positively correlated with independent measures of the same or similar domains (convergent validity). Conversely, there should be minimal associ- domains in question (divergent validity). In study C, domain scores were compared with blinded, independent clinician ratings of sexual functioning and with scales that measure marital adjustment (Locke-Wallace) and social desirability (Marlowe-Crowne). Significant positive correlations were observed between independent clinician ratings and subscale scores for all five domains (Table VI). In contrast, none of the correlations between domain scores and measures of marital adjustment or social desirability reached statistical significance.

Sensitivity and Specificity

To evaluate the sensitivity of the IIEF, a comparison was made between mean pretreatment and post-treatment domain scores of patients who were self-rated as treatment responders in study A. Specificity was assessed by comparing the pretreatment and post-treatment domain scores in patients rated as nonresponders in the same study. Patients were defined as responders or nonresponders based on their response to the end-of-treatment global efficacy question. All five domains of the IIEF demonstrated a high degree of sensitivity and specificity to the effects of treatment (Table VII). Although the magnitude of change was greatest for the erectile function domain, significant changes were observed across all five domains in the treatment responder group. The lowest magnitude of change was noted for the sexual desire domain. In contrast, none of the comparisons in the treatment nonresponder group approached significance (P

TABLE III. *IIEF domain scoring and intercorrelations*

Domain	Items	Score Range	Minimum Score	Maximum Score
		Domain Scoring		
EF	1, 2, 3, 4, 5, 15	0 (or 1)–5	1	30
OF	9, 10	0–5	0	10
SD	11, 12	1–5	2	10
IS	6, 7, 8	0–5	0	15
OS	13, 14	1–5	2	10

	EF	OF	SD	IS	OS
	Domain Intercorrelations				
EF	1.00				
OF	0.55	1.00			
SD	0.30	0.39	1.00		
IS	0.76	0.47	0.35	1.00	
OS	0.60	0.53	0.37	0.53	1.00

KEY: EF = erectile function; IIEF = International Index of Erectile Function; IS = intercourse satisfaction; OF = orgasmic function; OS = overall satisfaction; SD = sexual desire.

COMMENT

A 15-item, self-administered questionnaire scale was developed for the assessment of erectile function. This instrument (the IIEF) was developed in several stages, including initial pretesting with selected patient groups and expert panel consultants, followed by an intensive linguistic validation process. Based on a principal components analysis with varimax rotation, five factors or response domains were identified: (1) erectile function, (2) orgasmic function, (3) sexual desire, (4) intercourse satisfaction, and (5) overall satisfaction. The highest degree of positive correlation was between erectile function and intercourse satisfaction (r = 0.76), with two items (items 7 and 8) showing positive loadings on both factors. This is not surprising because a primary outcome of ED for most patients is the inability to achieve satisfactory sexual intercourse.[1]

Psychometric validation of the final instrument was addressed in three major areas: (1) test reliability, (2) construct validity, and (3) treatment responsiveness. Adequate performance in each of these areas should be demonstrated before a new scale is accepted for general research or clinical use.[21–23] For the IIEF, analyses were performed in each of these areas in two separate samples of patients with ED and age-matched controls. Overall, the IIEF was shown to have strong internal consistency, measured in terms of both the total scale and individual domain scores, and adequate test-retest repeatability. Although some variation in the degree of internal consistency was noted between samples, all of the values obtained were greater [illegible] 0.90. Test-retest repeatability correlation coefficients ranged from 0.64 to 0.84, and all were highly significant.

Construct validity (that is, whether the instrument actually measures what it was designed to assess) is normally accomplished by experimental testing of a priori questions or hypotheses, such as: (1) Will the test reliably differentiate between clinical patients and age-matched controls? (discriminant validity); (2) Can a positive association be shown with alternative measures of the same construct or domains? (convergent validity); and (3) Are the results influenced by related, but conceptually independent, variables? (divergent validity). In the present study, adequate construct validity was established in each of these three areas. Discriminant validity was demonstrated by a comparison of baseline scores between patients and controls. In the larger sample (studies A and B), between-group differences were highly significant (P values $\leq$0.01) for all five domains. In the smaller sample (study C), differences between groups were significant (P values $\leq$0.01) for all domains, with the exception of sexual desire (P = 0.72). In this study, patients and controls were closely matched on sexual desire, perhaps reflecting a high level of sexual motivation in patients seeking treatment in a clinical trial of ED. Tests of convergent and divergent validity were similarly confirmatory. First, a significant positive association was shown with independent clinician ratings for each of the major response domains. As expected, the highest correlation was observed for the domain of erectile function (r = 0.75). This association might have been even higher, except [illegible]

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.